WORKING PAPER · NO. 2026-18

Never Enough: Dynamic Status Incentives in Organizations

Leonardo Bursztyn, Ewan Rawcliffe, and Hans-Joachim Voth

JANUARY 2026

THE UNIVERSITY OF CHICAGO
BECKER FRIEDMAN INSTITUTE FOR ECONOMICS

5757 S. University Ave.
Chicago, IL 60637
Main: 773.702.5599
bfi.uchicago.edu

Never Enough: Dynamic Status Incentives in Organizations
Leonardo Bursztyn, Ewan Rawcliffe, and Hans-Joachim Voth
January 2026
JEL No. D22, D91, M52, N44

ABSTRACT

We study the ability of a firm to elicit repeated effort from workers by creating a “rat race” of hierarchical status-based incentives. We examine performance using data on over 5,000 German air force pilots during World War II. Pilots’ effort is hard to monitor; motivation is key to success. Fighter pilot performance increases markedly as they approach eligibility for a medal before falling off upon receipt of the award. The same effort path repeats itself as the pilot nears the next higher-prestige medal. Status-conscious pilots also exert more effort when new medals are introduced. We show that medals serve as substitutes for other forms of status. Medal cachet declines over time as lower-ability pilots receive them, making the introduction of new medals desirable. These results suggest that a tiered, expanding system of status-based incentives can repeatedly leverage worker status concerns to extract effort.

Leonardo Bursztyn
University of Chicago
Department of Economics
and NBER
bursztyn@uchicago.edu

Ewan Rawcliffe
Harvard University
Department of Economics
ewanrawcliffe@g.harvard.edu

Hans-Joachim Voth
University of Zurich Department
of Economics
and UBS Center for Economics
in Society
joachimvoth@gmail.com

Never enough / However much I do / However big I ever feel / It's never enough / Whatever I do to make it real / It's never enough ... / One more time to kill / Whatever I do it's never enough... / It's never, it's never enough.

Never Enough. The Cure

A central problem for firms is to ensure that workers provide effort, especially when it is difficult to observe. Financial incentives to elicit effort, such as sales targets or performance bonuses, are known to be effective as a one-off scheme (Lazear, 2018). However, financial incentives are, by definition, costly for the firm. Decreasing marginal returns might make the scheme prohibitively costly if used continuously. Can firms do better? Consider non-financial rewards, such as medals, awards, and prizes, to recognize particular achievements. Such recognition exploits people's tendency to compare themselves to others and their desire for status (Perez-Truglia, 2020; Cullen and Perez-Truglia, 2022; Frey and Gallus, 2017). Unlike other types of incentives, status is inherently *relative*. This creates an opportunity for a dynamic incentive system to exploit—or engineer—a never-ending cycle of goal chasing, keeping incentives sharp at a low cost. In this paper, we investigate whether a firm can tap into the human tendency to engage in status-seeking behavior as a tool to *repeatedly* motivate workers—and top performers in particular.

Static theories of non-financial rewards suggest their effectiveness as incentives (Besley and Ghatak, 2008). Moldovanu et al. (2007) show theoretically that creating different status tiers for employees may increase overall effort if ability levels are sufficiently diverse. To suggest how firms can sustain effort over time through status-based rewards, we note an analogy with theories of conspicuous consumption. Many firms employ tiered product hierarchies to exploit status concerns: airlines, car rental companies, casinos, and credit card providers offer silver, gold, and platinum tiers (Bursztyn et al., 2018). Pesendorfer (1995) formalizes this logic in a model of fashion cycles where a monopolist repeatedly introduces new product vintages. High-status consumers distinguish themselves by adopting the latest vintage, while older vintages lose their signaling value as they diffuse to lower types.

Whether such strategies can sustain worker effort in practice remains unclear. Status-based incentives must maintain their bite over time, yet rewards may lose motivational power once received. As in Pesendorfer (1995), tiered incentives could keep workers on a status treadmill, but only if carefully calibrated.[1] Status-based incentives may prove effective only over short horizons or when alternative status sources are limited.

We evaluate an organization's ability to create a rat race, inducing repeated bursts of worker effort through status-based awards. We study aerial combat during WWII. This setting has several advantages: fighter combat involves life-or-death stakes, outputs are well-measured, and pilots have high autonomy in choosing their level of effort (Ager et al., 2022). The task of a pilot is clearly defined but complex in practice. As is common in military and other bureaucratic settings, there is limited leeway to offer instrumental benefits. Hence, the armed forces typically rely on status-based incentives, such as medals. Moreover, the top pilots—fighter aces—contribute a disproportionate share of total output, making their motivation a priority in the design of the incentive system.

The German armed forces (*Wehrmacht*) engineered a sophisticated system of status-based incentives—the Knight's Cross medals: "the most sought-after military decoration" (Obermaier, 1966, p.21). The first Knight's Cross was initially introduced as the top award at the start of the war, requiring pilots to hold two lower (and older) awards for exceptional bravery to qualify. When many soldiers had reached this new top tier, additional medals—even more exclusive variants of the Knight's Cross—were introduced. The *Wehrmacht* thus created a "ladder" of up to five medals available to fighting men during the conflict. Such a system has the potential to motivate. For this to be true, the new rewards must alter the behavior of potential recipients who, otherwise, had few opportunities to distinguish themselves further. It could also backfire, creating jadedness about the repeated devaluation of (formerly) top awards and the introduction of additional carrots. How these opposing psychological forces play out is an empirical question.

In Section 1, we begin by presenting and discussing our dataset, which extends the setting used in Ager et al. (2022). They analyze personal rivalries between pilots and the role of envy as a motivating force, examining the victory and exit rates of former squadron peers when a pilot receives a mention in dispatches. This is a highly prized yet fleeting form of recognition,

[1] The closest empirical paper to ours, Landers et al. (1996), examines adverse selection in a one-off law firm promotion, where billable hours competition creates a "rat race".

making it helpful for studying spillovers from rivalry. However, their unpredictability makes mentions a poor source of direct motivation. In contrast, in this paper, we demonstrate the *direct* effect of a policy engineered to elicit effort through a highly prestigious, status-based award distributed via a mix of objective criteria and discretion.

Our results show that a carefully calibrated hierarchy of medals can motivate workers, inducing them to chase each new award. We study the effects of the first medal—the lowest rung of the ladder—on performance in Section 2. The Knight's Cross system was based on a set of (informal, loosely enforced) quotas, which were periodically changed throughout the war. It is essential to avoid naively measuring the spike in performance that mechanically precedes the receipt of a performance-based award (Frey and Gallus, 2017). We therefore employ an event study design over time, relative to meeting these *medal quotas*, to investigate medal effects.

A natural hypothesis is that pilot effort should increase as they approach this quota. Scoring an aerial victory always required risk-taking. However, as pilots approached the informal quota level, an additional reward beckoned for every extra victory: the (next) medal. Critically, pilots knew how close they were to the next quota. This information altered their perceived risk–return trade-off. In addition to the expected benefits of each marginal aerial victory, proximity to the quota promised a discrete jump in fame and recognition.

A large body of research suggests that such thresholds can generate sharp increases in effort. In psychology, the goal-gradient hypothesis suggests that individuals increase their effort as they approach a reward (Hull, 1934; Heath et al., 1999; Kivetz et al., 2006). In behavioral economics, models of quasi-hyperbolic discounting predict that agents place greater weight on immediate rewards, thereby magnifying the appeal of effort just before crossing the threshold (O'Donoghue and Rabin, 2015). Finally, salience theory suggests that performance quotas may act as prominent anchors in decision-making, so marginal progress near the threshold is over-weighted relative to the risks (Bordalo et al., 2012). While these behavioral frameworks differ in emphasis, they all predict that effort should rise steeply as pilots approach the medal quota.

Consistent with this, we find that pilots begin to significantly increase their level of effort from the baseline around three months before reaching the medal quota. We document sharp increases in performance during this period. The magnitude of the total additional output induced for ace pilots is on the order of six victories. This effect is significant; the median pilot in our dataset (of pilots with at least one victory) scores just three victories in their entire career. Once pilots received the medal, most of them scaled back their efforts drastically, returning to their former performance rhythm. During this period of enhanced effort, we nonetheless find no evidence of increased pilot exits.

We develop a novel procedure to identify the treatment effects of medal anticipation. The empirical challenge is a form of selection bias; only pilots who score victories (have a positive flow) can move closer to the victories quota (a stock). We call this the mechanical effect. This logic applies to any quota defined in terms of the victory stock, not just the real medal quota. We address this challenge by comparing the true performance path over time relative to meeting the real quota with performance paths over time relative to meeting placebo quotas. This allows us to isolate the causal component of effort from the mechanical effect of the victory flow increasing the stock. We verify the validity of this procedure using a Monte Carlo experiment.

In a complementary analysis, we leverage a unique feature of our historical setting to facilitate the causal identification of this anticipation effect. The medal quotas required for each medal changed over time and location. When quotas changed, some people who were previously close to the threshold were now far from the "zone" of medal eligibility; others were suddenly closer. Such changes are exogenous to individual pilot performance. When we instrument "in zone" status with quota changes, we obtain similar results to our preferred event study procedure.

To understand what motivates pilots to exert extra effort in pursuit of medals, we examine alternative sources of prestige. Pilots already holding other medals unrelated to the Knight's cross award scheme, senior officers, and noblemen appear *less* motivated by the prospect of the medal than pilots without alternative sources of social status. This is in line with the prediction of status signaling models (see, e.g. Spence, 1973)—pilots with strong existing signals have weaker incentives to invest in costly additional signals. This pattern demonstrates that medals act as substitutes for other status signals, highlighting the importance of status competition as a mechanism. Moreover, it suggests that patriotic or ideological factors are unlikely to be the primary drivers of our results, thereby enhancing external validity.

The second unique feature of our setting allows us to analyze the emergence and perpetuation of a "rat race" in Section 4. New medals were repeatedly introduced. The first medal was available from the start of the war; four additional medals were subsequently created. We use the staggered introduction to demonstrate that the top pilots prize higher-tier medals as they ascend the ladder. The quota criteria for the higher medals are strictly and often significantly greater than those for the preceding medals. Moreover, regulations required that the medals be awarded in sequence.

Using a generalized synthetic control approach (Xu, 2017), we show that pilots who are medal sensitive – reacting strongly to being close to the medal quota – increase their efforts immediately following the *introduction* of the new medals, compared to other pilots who are less sensitive to the prospect of medals. This pattern holds for the introduction of the second medal in June 1940 and the (simultaneous) third and fourth introductions in July 1941. In other words, the expanding hierarchy of medals captured the imagination of high performers, even if they had already achieved substantial success.

In addition, pilots also repeat—almost exactly—the effort increase seen for the first medal quota as they approach the new, second (and subsequent) medal quotas. We show a double event study, where event time is defined relative to the treatment of sequentially meeting the first and second medal quotas. "Making it"—once—for these pilots is insufficient to nullify their desire for additional distinction. In combination, these results suggest that the creation and perpetuation of a system of status incentives were broadly effective for the Luftwaffe.

To rationalize the efficacy of the rat race engineered by the German air force, we exploit insights from models of fashion cycles in Pesendorfer (1995). For status goods, consumption is a signal that loses value over time as more consumers purchase the good. Eventually, high-type consumers demand a new good that will allow them to send a distinguishing signal. The firm finds it profit-maximizing to produce this novelty. Central to this conclusion is that, beyond the general expansion in the good's supply over time, high-type consumers buy the product early, and low-type consumers buy later. We test for an analogous feature in our data by examining the ability rank of pilots who receive the medal over time. We measure ability as the permanent productivity of pilots, purged of peer spillovers, given the team environment (Mas and Moretti, 2009). The ability rank of pilots awarded any given medal declines over time. Therefore, each recipient is faced with a gradual 'cheapening' of the award he holds, as lower-ranked pilots also receive it, reducing the status implicitly attached to the award. In line with the fashion cycle model, this motivates both the dynamic introduction of a new medal and ensures that the new award induces effort from high-types.

Our paper contributes to the extensive literature on organizations using economic (Kamenica, 2012) and social incentives (Ashraf and Bandiera, 2018) to solve contracting problems. Building on research into standard monetary incentives and their effects on performance (Prendergast, 1999), a strand of work has focused on behavioral and symbolic forms of worker incentives (Kosfeld and Neckermann, 2011; Ashraf et al., 2014; Ager et al., 2022), establishing a role for social recognition and signaling. Akerlof and Kranton (2005) discuss the role of identity, particularly in military settings, as a source of motivation. Our focus is specifically on a richer setting where the firm cares about the dynamic problem of extracting repeated effort. One prominent form of recognition, awards, is studied in Frey (2007) and Frey and Gallus (2017). The latter study highlights the empirical difficulty of examining the effect of performance-rewarding awards on performance, a challenge we overcome by exploiting the quota criteria that guide the supply of medals.

In tournament theory, the classic work of Lazear and Rosen (1981) examines optimal rank-order compensation schemes for risk-averse agents heterogeneous in ability. They describe how non-linear compensation schemes, such as promotions (or medals in our present study), induce effort and risk-taking due to tournament competition among workers. However, their model describes a case where contemporaneous output is generally costly to observe. In related work, Besley and Ghatak (2008) directly models status incentives, including medals and job titles. They show that awards can induce effort under zero marginal cost to the principal in a static setting, even if effort is non-verifiable but output is observed. We empirically demonstrate the feasibility of such systems in a dynamic setting. Recent work has incorporated behavioral preferences into tournament design. Gill and Stone (2010) and Eisenkopf and Teyssier (2013) show how relative rewards, fairness consideration, or loss aversion can affect incentives and effort.

Our study's central form of incentive is a status symbol, building on work studying the economics of status. On the consumer side, Bursztyn et al. (2018) identifies significant demand for a higher-tier good with otherwise identical utility to a lower-status alternative. A related mechanism is explored by Imas and Madarasz (2022), where the possibility of excluding others from consuming a good increases desirability. The Knight's Cross medals are likewise a tiered system of signals, with the standards for receiving the award altered over time to restrict supply. Moldovanu et al. (2007) and Rayo (2013) examine the theoretically optimal behavior of organizations that seek to profit from workers' status concerns by assigning status categories or selling status signals, but in static settings. Our data matches two key predictions of Moldovanu et al. (2007). First, despite the zero-sum nature of status competition, several coarse status categories are optimal under concave distributions of ability. Second, the very top performers should not be pooled with lower types. Hence, we provide empirical evidence for how such a scheme works and continues to work over time. A more distant relation is found in Ray and Robson (2012), which demonstrates the endogenous emergence of risk-taking behavior when incorporating status concerns into an otherwise standard growth model.

We speak to a particular literature on dynamic incentives that documents cyclical patterns of effort. Pesendorfer (1995) is a central reference to explain how the dynamics of a status signaling game evolve between consumers and a monopolist. We see a powerful analogy to this process in the organizational behavior we study, whereby the increasing proliferation over time of a status good motivates the innovation of a new good to be consumed by high types. We provide empirical evidence that a similar process can guide a system of status-based incentives. Misra and Nair (2011) studies the optimal design of a compensation scheme with a quarterly sales quota. They document time paths of output as a function of distance to the quota with distinctive spikes right below the quota, motivated by fears of a ratcheting effect if quotas are far exceeded. We present comparable results showing that workers do not smooth their effort over time in response to discrete, quota-driven incentives. Genakos and Pagliero (2012) examine competitive weightlifters and show that highly ranked individuals perform less well; revealing relative rank leads to greater risk-taking.

Our rat race is related to the concept of hedonic adaptation and goal-setting. The idea of a rat race was formalized by Akerlof (1976), in which the "rat" (worker) can benefit from working harder not just by producing more output but (perhaps centrally) by signaling his superior type, which otherwise cannot be observed. Landers et al. (1996) also studied a rat race and showed empirically that the adverse selection described by Akerlof induces overworking in US law firms. However, their setting does not allow for consideration of the dynamics of the rat race, which we can study through the repeated pursuit of new medals. Hedonic adaptation (Brickman and Campbell, 1971; Frederick and Loewenstein, 1999) has been studied in economics by Bottan and Perez Truglia (2011) and Perez-Truglia (2012), who document the habituation of happiness following major life events. More generally, the role of goal-setting in guiding effort and performance is summarized in Genicot and Ray (2020).

Allen et al. (2016) document bunching in performance around reference marathon times, even when these benchmarks lack external incentives. In Bénabou and Tirole (2011), a treadmill effect describes how accumulated status can lead to a "self-defeating" over-acquisition of yet more status. In our setting, the expanding range of status symbols amplifies such effects.

The paper proceeds as follows. Section 1 provides background on the setting and our data sources. We analyze the main effects of medal chasing in Section 2. Section 3 analyzes the organization's problem and presents results on medal cheapening. Section 4 presents our results on the medal ladder before we study mechanisms and overall magnitudes in Section 5. We conclude in Section 6.

1 Setting and Data

In this section, we briefly review the role and performance of the German air force (*Luftwaffe*) in World War II and describe the system of Knight's Cross medals. We then summarize our data sources and construction.

1.1 The German air force

After WWI, the Versailles Treaty abolished Germany's air force. After 1933, the Nazi regime nonetheless began to build a new air force under the energetic leadership of Hermann Goering, a former fighter ace and squadron peer of the Baron von Richthofen. The *Luftwaffe* was mainly designed to provide close air support to the army. To this end, it needed to attain air superiority over the battlefield, leading to a focus on fighter and ground attack aircraft. The German air force invested heavily in advanced aircraft, such as the monoplane fighter Bf-109 and the dive-bomber Stuka, and developed new tactics, such as the four-finger formation. Both equipment and tactics were tested and refined during the Legion Condor's service in the Spanish Civil War, the German support formation for General Franco's putsch.

When WWII broke out, the campaigns in Poland and France quickly showed the merits of close air support and the importance of air supremacy. However, the German air force notably failed to subdue the RAF during the Battle of Britain, having met its match in men and material (Bungay, 2010). During the opening phase of the campaign in Russia, the Luftwaffe dominated the air war, destroying much of the Soviet air force on the ground. The Red Air Force eventually recovered, outnumbering German forces from 1942 onward. Still, it never matched the quality of either pilots or machines one-for-one. On the Western Front and over the Reich itself, air superiority was lost by 1944, when long-range fighters escorted Allied bombers. Despite new aircraft such as the first jet fighter, the Luftwaffe was no more than a shadow of its former self after the spring of 1944. As Allied bombing destroyed fuel plants, training suffered; pilots were dying at a rate of 25% or more per month by mid-1944 (Murray, 2015).

The campaigns in France and Britain had seen a small, select group of pilots attain *Experten* (ace) status, having shot down 15 or more enemy planes. Because of the one-sided air war in Russia, total victory claims for the Luftwaffe surged after 1941. German pilots were never rotated out (in contrast to standard practice among the Western allies), even after extended tours ("fly till you die"). Therefore, top German overall tallies exceed those of Allied pilots by a large margin. While testament to the pilots' skill, detailed comparisons looking at the length of combat and the number of enemy planes normally fails to show any major advantage of the best German over the best Allied pilots[2].

Victories mattered greatly for a pilot's status among his peers, in the air force, and in society at large.[3] Air supremacy was a vital component of *Blitzkrieg*. The air force High Command dedicated substantial resources to tracking and verifying the aerial victory claims of individual pilots. Gun camera footage was used when available. All aerial victories were entered into a pilot's logbook with information on the type of the enemy plane, how it was shot down, whether the impact was observed or not, and in which grid coordinates the wreckage went down; they had to be witnessed by a fellow pilot or confirmed by a corresponding ground unit (for example, by documenting wreckage). While over-claiming no doubt occurred, the consensus in the literature is that German victory records are on the whole reliable, erring consistently on the side of caution (Frieser, 2007).[4]

[2] Simkin and Roychowdhury (2006) examine this question for German pilots during WWI, when similar incentives obtained, and find no sign of outperformance.

[3] Propaganda regularly highlighted the accomplishments of German pilots. Pilots pictures were distributed as postcards, similar to portrait pictures of movie stars.

[4] Total German victory claims are typically less than the sum of Allied aircraft destroyed, according to Allied records, which suggests underclaiming.

1.2 Knight's Cross medals

The Knight's Cross (*Ritterkreuz*, KCR) medals were the "most sought-after military decoration" (Obermaier, 1966, p.21) awarded to German soldiers during WWII. Initially a single medal, the Knight's Cross of the Iron Cross was introduced in September 1939 at the outbreak of the war. It was an all-purpose medal that was awarded with a particular emphasis on individual bravery (Williamson and Bujeiro, 2004, p. 3).[5] The Knight's Cross was part of a "ladder" of awards—only recipients of the Iron Cross 1st class and 2nd class qualified. Introduced during the War of Liberation in the early 19th century, the Iron Cross was deliberately conceived as a meritocratic award, bestowed without regard for rank or noble origin. Iron Crosses were awarded during the 1870/71 war against France, during WWI, and WWII.

As the war continued, numerous high-performing soldiers such as successful U-boat captains or tank "aces" received the KCR. The armed forces soon began to argue for the introduction of additional medals to recognize the continued performance of outstanding fighting men. The Luftwaffe was particularly vocal in lobbying for new awards. To satisfy this need, higher-ranked variants of the Knight's Cross were introduced, starting with the Knight's Cross of the Iron Cross with Oak Leaves in 1940, followed by the Knight's Cross with Oak Leaves and Swords, the Knight's Cross with Oak Leaves, Swords, and Diamonds in 1941, and the Knight's Cross with Golden Oak Leaves, Swords, and Diamonds in 1944.

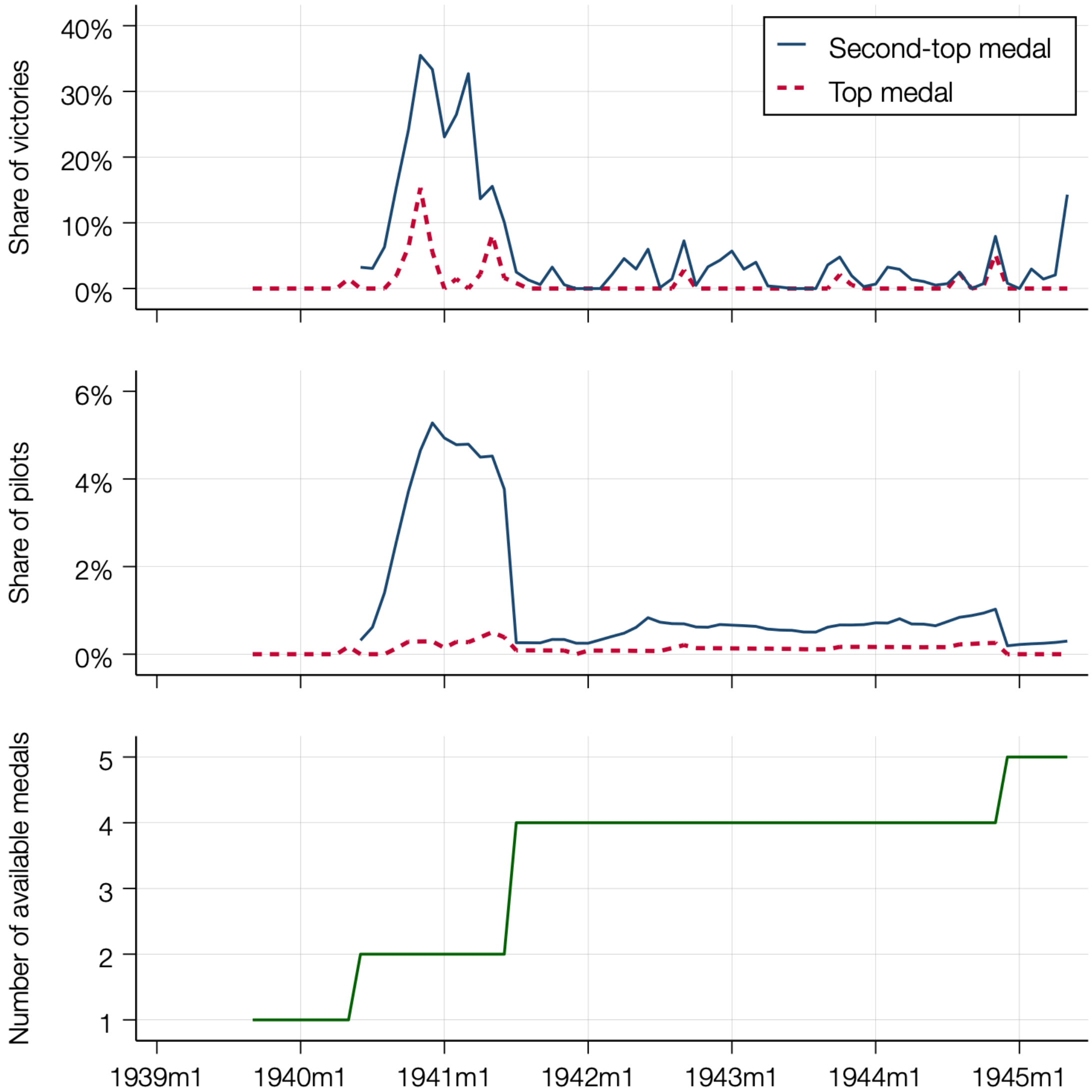
Knight's Cross medals were worn around the neck, making them a particularly prominent part of a soldier's uniform. Remarkably, they partly reversed the military hierarchy of deference and respect—even high-ranking officers had to salute decorated privates or sergeants. Only 7,175 were ever awarded, including 882 Oak Leaves, 159 Swords, 27 Diamonds, and one Golden Oak Leaves. While most awards went to the German army, the air force had a high concentration of awardees, given the number of eligible personnel. A total of 569 fighter pilots received the Knight's Cross.

Figure 1 provides an overview of the Knight's Cross system. The lower panel shows the number of Knight's Cross medals available to pilots across time, highlighting the staggered rollout of additional medals. The middle panel shows the share of pilots holding top and second-top medals over time. In our data, we see that close to one in twenty pilots gets

[5] While awards to high-ranking officers occurred, the emphasis on personal bravery increased sharply during the war. From 1942 onwards, less than half of the medals awarded went to officers, compared to around four-fifths in 1940.

close to "topping out" of the incentive ladder. The connection to the supply of medals in the lower panel suggests that the increasing concentration of medalists preceded new awards introductions. The upper panel shows the share of all aerial victories achieved in each month by top medal holders, which is highly disproportionate to their share in all pilots. Controlling the incentives of medal-type pilots was therefore crucial. We henceforth refer to the Knight's Cross system of medals by medal 1, medal 2, etc., where medal 1 $<$ medal 2 $<$ medal 3.

Figure 1: Introduction of new medals



Notes: Bottom panel shows the number of medals available to pilots as a function of time. The medals, in order of introduction, are: Knight's Cross of the Iron Cross (*Ritterkreuz*), Knight's Cross of the Iron Cross with Oak Leaves (*Eichenlaub zum Ritterkreuz*), Knight's Cross with Oak Leaves and Swords (*Eichenlaub mit Schwertern zum Ritterkreuz*), Knight's Cross with Oak Leaves and Swords and Diamonds (*Eichenlaub mit Schwertern und Brillanten zum Ritterkreuz*), Knight's Cross with Golden Oak Leaves, Swords and Diamonds (*Goldenes Eichenlaub mit Schwertern und Brillanten zum Ritterkreuz*). Middle panel shows the share of pilots in each period who have achieved the highest and second-highest Knight's Cross medals available at that time. Top panel shows the share of victories in each month attributed to pilots holding (at least) the indicated medal.

To allocate medals, the Luftwaffe High Command created a cumulative aerial victories target that defined when a pilot was eligible to be nominated for a Knight's Cross. These targets were informal; they did not constitute a strict "quota". Nonetheless, quota guidance was "brought to the attention of field commanders." Discretion was involved in every nomination, which needed to be passed up and endorsed by every level of the chain of command before being finally accepted (and often personally awarded) by Hitler himself. Pilots themselves were keenly aware of quota targets. Their aircrafts' tails displayed their "tally", keeping everyone aware of how they and their peers had performed. Obermaier (1966) argues that the quota system was explicitly created as "psychological bait".

The initial quota for medal 1 was set at 20 victories, a threshold reached by just six pilots in the first twelve months of the war. However, as the war went on, this quota was no longer appropriate to maintain the level of prestige desired by the High Command. On the Eastern front, where victories initially came easily against Soviet pilots with limited experience and inferior equipment, the quota was met especially quickly. To maintain scarcity, the quota was adjusted upwards, and a new quota was introduced for pilots on the Eastern front. As the war turned against the Germans, the quota was adjusted downwards. Medals 2 and 3 also had unique quota requirements. The specific levels of each quota are taken from the historical account in Obermaier (1966, p.p. 21–23) and shown in Figure A1. We see substantial variation over time in the level, reflecting the desire of the High Command to preserve a consistent standard amid changing combat conditions. The quota requirement for medal 3 at one point rose to 250 victories, an astounding target given that the median pilot in our dataset (which is already restricted to pilots who scored at least a single victory) finished the war with just three victories.

The medal system could be modified as priorities changed. Starting in 1943, Allied bombing raids became a growing threat. Bomber aircraft, especially the USAAF B-17 "Flying Fortresses" were a daunting opponent. The difficulty of shooting down a bomber was clearly greater than of that of destroying a single-engine fighter: Equipped with up to 13 .5 inch machine guns and increasingly accompanied by a fighter escort, Allied bombers accounted for a high share of German aircraft losses. To align incentives, a points system (*Punktwertung*) was introduced that differentially rewarded the pilots based on the type of adversary and their role in their destruction. However, there are reasons to be skeptical of the importance of the points system over the quota system. Obermaier (1966, p. 21) highlight resistance toward the new system from pilots, who disliked the distraction from the purity of fighter versus fighter combat, and the complications of group effort necessary to attack bomber formations (with the associated challenges of awarding credit for a "kill"[6]). Second, empirically, we find no evidence that the predictive power of the victories quota in the award of medals falls after the *de jure* introduction of the points system in October 1943. We, therefore, use the victories quota system as the salient target for pilots throughout the analysis.

The dispatch mentions studied in Ager et al. (2022) and the Knight's Cross medals are complementary incentives used by the German air force during WWII. Both were designed to encourage and reward exceptional performance. Awards and mentions are correlated, but only weakly ($\rho = 0.1567$). Three main differences make medals particularly useful for our purposes. First, medals were permanent and visible, in contrast to the fleeting nature of the recognition provided by a mention. Recipients wore their medals prominently around their necks on all occasions, including while flying (Williamson and Bujeiro, 2004). Figure A2 shows the appearance of medal 1. Even NCOs had to be saluted by superior officers when wearing the Knight's Cross. Second, the award was not only highly prestigious—as underscored by many awards made by Hitler himself. It was also more predictable and "attainable"—decisions to award the medal were based on the informal quota system, making it more fair and transparent than mentions in dispatches. There were approximately six times more Knight's Cross medals awarded in the course of the war than there were mentions in dispatches.[7] Finally, the medal system constituted an explicit hierarchy of awards, with new, higher levels added throughout the war. A medal was only available to pilots who held the proceeding medal. For all these reasons, the Knight's Cross system is helpful since it allows us to study the rat race of pilot performance created by the award system.

1.3 Data

Performance and ability

Our data on the deployment and performance of German fighter pilots is described in detail by Ager et al. (2022), who construct a detailed panel by combining several comprehensive secondary sources based on microfilm data from the German Federal Archives. The data

[6] The air force tried to address some of these issues for awarding points for forcing bombers out of formation, in addition to destroying it.

[7] Mentions in dispatches do not always include first names, complicating identification; the range for mentions across the whole armed forces is 1,182-1,739, while around 7,300 Knight's Cross medals were awarded. For fighter pilots in particular, our data has 85 mentions and 558 Knight's Cross medals.

covers daytime fighter pilots (engaged in air-to-air combat) in the Luftwaffe during World War II from September 1939 to May 1945. The sample is limited to pilots who score at least one aerial victory during this time. Summary statistics for the pilot–month panel (N $= 82,348$) are in Panel A of Table A1, with pilot-level characteristics given in Panel B (N pilots $= 5,081$).

Our primary outcome in this panel is the monthly victory rate: counts of enemy aircraft shot down and verified by High Command. Additionally, we study pilot deaths in combat using a dummy that codes exit from the panel before the end of the war. Around half of the exits are documented in two separate sources. Two-thirds of exits represent the death of the pilot, while the remainder represent their being shot down and/or seriously wounded, such that they do not continue flying.

Pilot ability is an obvious confound in any correlation between incentives and performance. Simultaneously, peer ability confounds individual performance since squadron mates play an important tactical role, even when not directly achieving aerial victories. An ideal variable would be an ability score assessed in an objective way based on pilot performance under repeated random peer assignment. We follow the estimation procedure in Mas and Moretti (2009) to give workers' individual permanent productivity in a setting of individual production with peer spillovers. To adjust for the large fraction of unique peer combinations relative to observations in our data, we use the network reduction procedure from Hamilton et al. (2022). This allows us to recover an estimate of time-invariant, inherent ability for most of our pilots. We interpret the residual of performance on ability and other observable fixed effects (period, combat theatre, squadron) as effort. Appendix B provides greater detail on this procedure. We use the words 'ability' and 'quality' interchangeably throughout.

Medals

We collect data on the identity of medalists and the timing of awards from *Lexikon der Wehrmacht* (`https://www.lexikon-der-wehrmacht.de/`), a website housing extensive databases on German military history. The site provides the full name, military branch, unit, and rank of each medalist, alongside the date of award and type of medal. We manually harmonize the spellings of names, ranks, and units and then merge the medal database into our panel. We perform the same exercise to verify the data with an alternative source, the Kracker Luftwaffe Archive (https://aircrewremembered.com/KrackerDatabase/). We find the data to be highly similar but prefer the *Lexikon* data for its completeness. Panel B of Table A1 shows that around 8% of pilots achieved medal 1, 2% medal 2, and 1% medal 3. The number of daytime fighter pilots who received medals 4 (6) and 5 (0[8]), the highest medal, is too low for our analysis, despite their creation and existence being important to our mechanisms.

We collect medal quotas from Obermaier (1966). The source is narrative rather than tabular, so assumptions must be made to complete a month-by-front-by-medal panel of the quota level. Our preferred interpretation was hand-coded by a research assistant without knowledge of our research design, reading from the German text. We then compared this interpretation to the results of three other coders, who worked from an English translation (the source provides both languages). While there are minor differences between coders, the discrepancies are small and do not have a substantive impact on any of our analyses. We show robustness to an alternative coding of the quota level in the Appendix.

2 Medal-Chasing

In this section, we examine whether pilots in general 'tried harder' and had greater success as the approached the quota for the first top-tier medal, the Knight's Cross. First, we describe the non-standard features of the econometric problem posed by this problem. Second, we describe our proposed estimator, which uses the victories quota and a mechanical bias correction generated using placebo quotas, which we validate using Monte Carlo simulations. Finally, we present our estimates after bias-correction.

2.1 Empirical design

Do pilots respond to the prospect of receiving the first medal in the Knight's Cross hierarchy? We aim to measure the effort induced by the anticipation of eligibility for a medal. In Figure 2, we plot the mean victories of pilots who are about to receive the Knight's Cross (in blue), and compare their performance with peers at the same airbase and calendar time (in

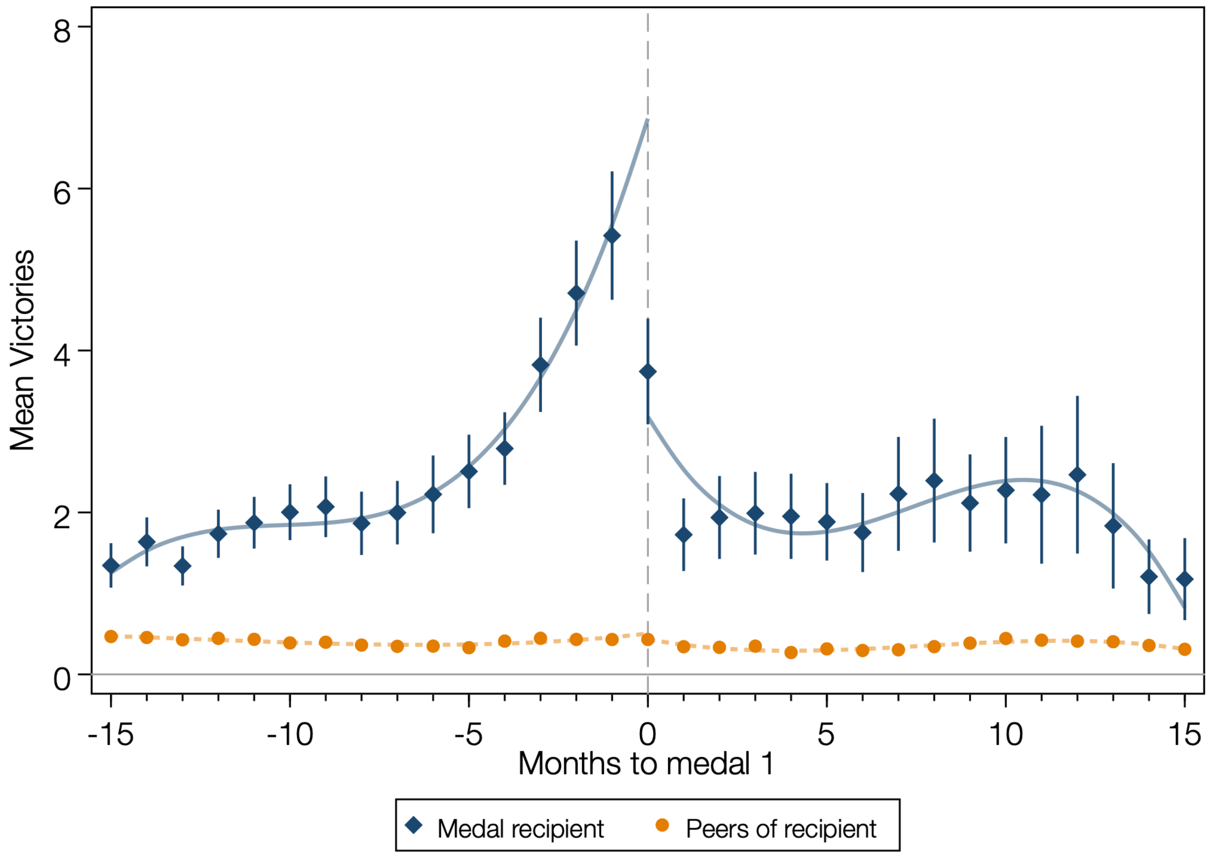
[8] The highest medal was only awarded to a ground-attack pilot, Hans-Ulrich Rudel.

orange). The raw data show a suggestive spike prior to medal receipt, before a sharp decline in performance—but the exercise in Figure 2 conditions on the outcome (medal receipt).

Our goal is to measure how much of the increase in effort prior to the medal is driven by additional effort. We do so in two steps. First, we use the medal *quotas* instead of actual medal receipt to define event time. Pilots did not know if and when they would receive the medal—but they did know how far from the relevant quota they were. Second, we cleanse the increase in performance of upward bias, arising from the fact that as a pilot scores victories, they mechanically approach the target value.

The medal quota has strong predictive power for when pilots are awarded medals. We examine the likelihood of receiving medal 1 as a function of distance to quota, defined as current cumulative victories minus the quota level in Figure 3. The likelihood of an award ('medal risk') peaks just after pilots reach the quota, consistent with flexible, yet relevant, eligibility criteria. Indeed, the fact that the quotas are not perfectly binding was an explicit part of the system. The slight offset of the peak risk from a quota distance of zero also likely reflects procedural delays as the medal nomination is passed up the chain of command. This process is shown in Figure A3. In the language of Frey and Gallus (2017), Knight's Cross medals are confirmatory rather than discretionary awards that correspond to known criteria. With this in hand, we can continue our analysis, considering event time relative to meeting the medal quota—the timeline that pilots can anticipate and control.

Figure 2: Raw victory rates over medal event time



Notes: Figure shows mean monthly victories $\bar{Y}_j$ in blue over time relative to winning the medal 1. The orange line shows the mean monthly victories of peer pilots who fly at the same airbase and in the same month as the focal pilot (in blue). Bars give the 95% confidence intervals for the mean. Lines are splines fitted to cubic polynomials of the data.

Do pilots closer to the quota for a medal increase effort? We hypothesize an effort profile as a function of time relative to (likely) medal receipt with two key features. First, the medal quota (the target for the *stock* of victories to be attained in order to receive the medal) allows pilots to dynamically anticipate the likelihood of receiving a medal. Pilots should increase their effort (*flow* of victories) as they approach the quota. We will interpret the sum of the effort effects induced by the quota in the period before the pilot meets the quota as an "in zone" effect. If large, this effect measures the extent to which pilots are motivated by the medal. Second, effort should fall sharply after the quota is exceeded, since pilots have received or will shortly receive the medal. This feature helps to confirm that the effort profile is indeed related to the medal and not otherwise spurious. Consider the empirical model:

$$\bar{Y}_j = \theta_j + \varepsilon_j \tag{1}$$

Define the event time E_i as the period t in which pilot i meets the medal quota[9]. $\bar{Y}_j$ is the mean victory rate of pilots $E_i - t = j$ months from meeting the medal quota, θ_j is the effort effect we aim to recover and ε_j is the error term.

Figure 3: Medals are awarded according to an informal victories quota

Notes: Figure shows the relationship between distance to the informal victories quota and the risk of being awarded medal 1. The horizontal axis shows the difference between a pilot's cumulative victories and the informal victories quota for the medal, which varies by period and front. The red line gives the local polynomial risk of being awarded medal 1 conditional on distance, bounded by 95% confidence intervals in the shaded area.

The econometric challenge in recovering θ_j is that we expect non-zero ε_j. In particular, we expect systematically positive errors due to two distinct selection effects. The first reason is

[9] Medal quotas varied by combat front, and switching front significantly affected performance independently of the medal system. Therefore, we do not allow a pilot to be defined as meeting the quota immediately after they switch front.

mechanical selection: positive errors mechanically advance the event time index. Each additional victory today moves you closer to meeting the medal quota in the future. Informally, this is the challenge of regressing a flow on a stock. More strongly, under a null hypothesis of $\forall_j : \theta_j = 0$, we still expect $\bar{Y}_j$ to be large just before the medal is achieved because only positive shocks can move you towards the quota. At the very least, there will be an average error of one in the treatment period, since this size of shock is sufficient to push a pilot over the quota threshold. This is related to Ashenfelter's Dip, when low earnings mechanically drive selection into a worker training program, thereby biasing estimates of post-treatment earnings effects that leverage comparisons to the dip (Ashenfelter, 1978). Our problem, nonetheless, is distinct from this literature: we seek to estimate pre-treatment effects, so our design is fundamentally not a difference-in-differences design.

The second reason is a *survivorship* effect. Latent pilot ability drives both the victory rate and survival. These components are highly correlated: pilots are at a higher risk of dying in a combat-intensive environment. Pilots who meet the quota are positively selected for ability, as one would expect. This selection is apparent in the time-invariant level difference between the blue and orange lines in Figure 2, indicating that the medalists are always performing more highly than their peers. A standard fixed-effects adjustment can address this level difference. However, further correction is needed to account for changes in composition as the event-time index advances, because higher-ability pilots also tend to survive longer.

Given the event study framework we have described, we posit a linear panel model, adjusting for nuisance ε_j terms, which we refer to as the "mechanical effect". The standard regression equation is:

$$Y_{i,t} = \alpha_i + \delta_t + \lambda_f + \varphi_i + \sum_{j \in J} \mu_j \mathbb{1}\{t - E_i = j\} + Y_{i,t-1} + \epsilon_{i,t} \tag{2}$$

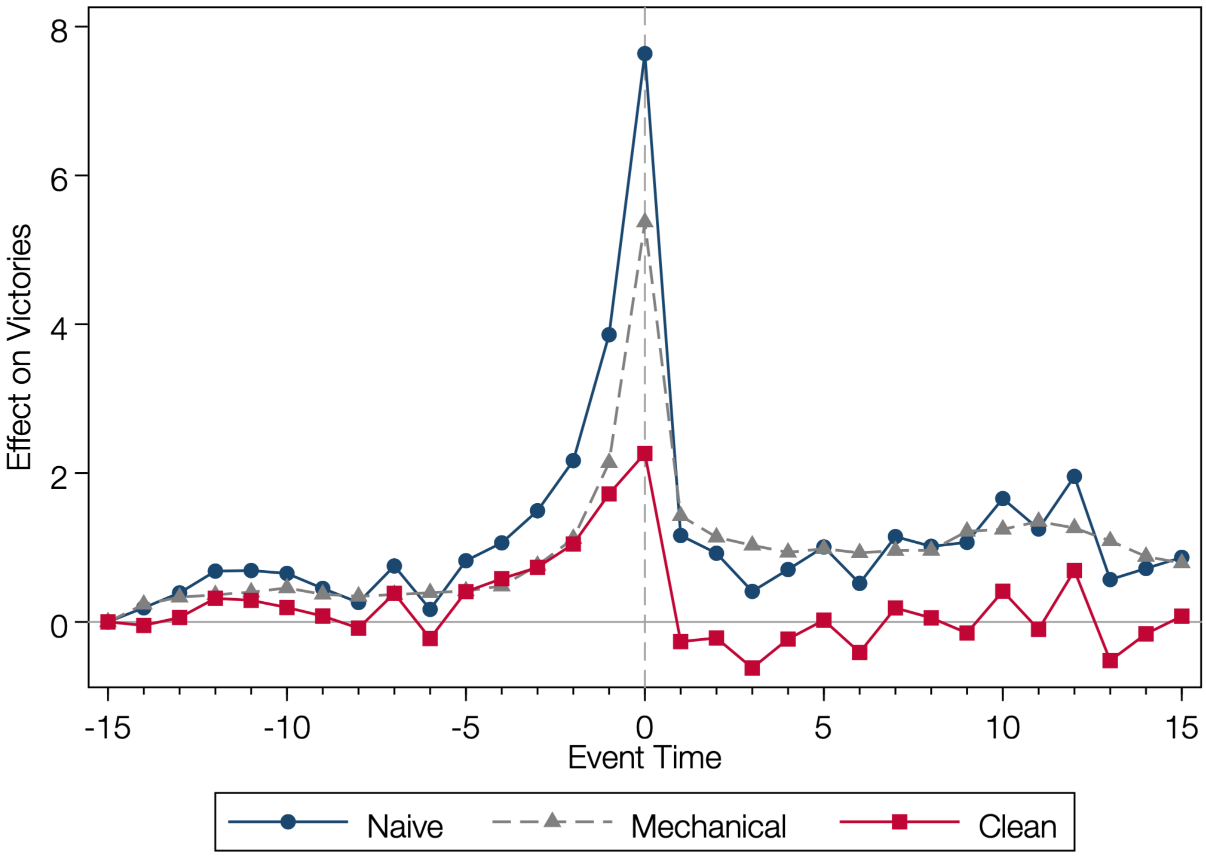
We denote unit and time-fixed effects with α_i and δ_t, respectively. The dummy variables λ_f and φ_i adjust for differences in combat environment across the Eastern and Western fronts and for holding the medal, respectively. We include an auto-regressive term, $Y_{i,t-1}$, in line with the literature on dynamic panel count models (Cameron and Trivedi, 2015), and explicitly consider a Poisson count model as a robustness exercise. The regression residual error is denoted by $\epsilon_{i,t}$.

We fully saturate the model shown in (2) with treatment lags and leads, with a leave-out period of $j = -15$ and with lags and leads $|j| \geq 15$ accumulated into end bins. The choice of the leave-out period is explicitly motivated by anticipation effects in periods closer to $j = 0$. However, we still expect that the μ_j terms are contaminated with positive bias from the mechanical effect.

Our solution is to exploit the fact that the mechanical effect can be recovered by studying quotas for which there is no associated treatment effect by construction: placebo quotas.[10] As a pilot approaches a placebo quota, their effort is unchanged, but their mechanical effect will be recovered by estimating (2). We can then subtract out the mechanical effects under the placebo quota event study from the true event study coefficients to recover θ_j. Specifically, we generate many placebo quotas and create a stacked dataset containing both the true lags and leads of the event time, as well as the placebo lags and leads. The clean effect is then estimated as the event study coefficients interacted with a variable denoting the true lags and leads. We show in Figure 4 the separation of the mechanical effect ε_j from the 'clean' causal effects θ_j. The naive estimates correspond to the μ_j in (2) over the real data. We estimate the clean effects using our procedure, drawing 25 placebo quotas per pilot, and compute the mechanical effect as the residual difference between these estimates. As expected, the mechanical effect is i) everywhere positive and ii) strongly peaking just before the quota.

[10]This is in the spirit of Borusyak and Hull (2023), who 'recenter' estimates after specifying a model of nonrandom exposure to shocks.

Figure 4: Decomposition of the event study into mechanical and clean effects



Notes: Figure shows the decomposition of coefficients from an event study over time to meeting the quota for medal 1 under our mechanical-correction procedure. Th *Naive* series plots the coefficients from a 'naive' estimation of (2) on 82,348 observations without any adjustment. The *Clean* series shows the coefficients recovered after the mechanical-correction procedure that correspond to θ_j with $p = 25$ placebo quota draws per pilot. The *Mechanical* series is the difference between these two estimates and corresponds to the error term ε_j.

This approach relies on several assumptions for identification. First, we assume that the leave-out period, $j = -15$, is sufficiently distant from the quota that pilots have not yet adjusted their effort in response to it. This does not preclude that a pilot's baseline level of effort is elevated by the existence of the medal system — such an effect is partialled out by pilot fixed effects. Second, we assume that, averaged across draws, the placebo quotas share a common error structure with the real quotas: $E[\varepsilon_j^{\text{placebo}}] = \varepsilon_j^{\text{real}}$. We aid the plausibility of this assumption by drawing placebo quotas from a distribution that is empirically similar to the range of the real quotas[11]. Third, we assume that other determinants of performance

[11] We draw placebo quotas uniformly from $q_{it} \cdot [\underline{\mu}, \overline{\mu}]$, where q_{it} is the real quota facing a given pilot. We set $\underline{\mu} = 0.1$ and

(e.g, tactical reassignments, weather, materiel) are not systematically timed relative to the quota. While rich fixed effects can provide conditional exogeneity with respect to some of these factors, we also show in Figure A5 that covariates that capture a range of these timing concerns do not systematically coincide with the quota event.

To verify that this approach recovers the effort effects θ_j, we perform a Monte Carlo experiment. Our DGP for synthetic data broadly matches the size and empirical features of our pilot-month panel and includes endogenous exit driven by enhanced risk-taking and randomness. The full specification is given in Appendix C. We generate a medal anticipation effect that is a modified logistic function conforming to our alternative hypothesis of a positive in zone effect. We also test the estimator under a null hypothesis of $\forall_j : \theta_j = 0$. Figure A6 visualizes the results of 50 Monte Carlo simulations, where within each simulation we generate $p = 25$ placebo quotas per pilot. The dashed gray line shows the raw estimator corresponding to (2): it overshoots the true pre-quota effort by roughly 40 percent and exhibits a spurious positive effect *after* the quota has been reached. By contrast, the solid red line (clean estimates) tracks the black true effect almost perfectly and its confidence band covers θ_j across the event window, under both the null and alternative hypotheses. The simulation, therefore, confirms that the adjusted estimator removes the mechanical bias. In all the individual-level results to follow, we apply this mechanical-correction procedure to our estimates.[12]

2.2 Individual performance results

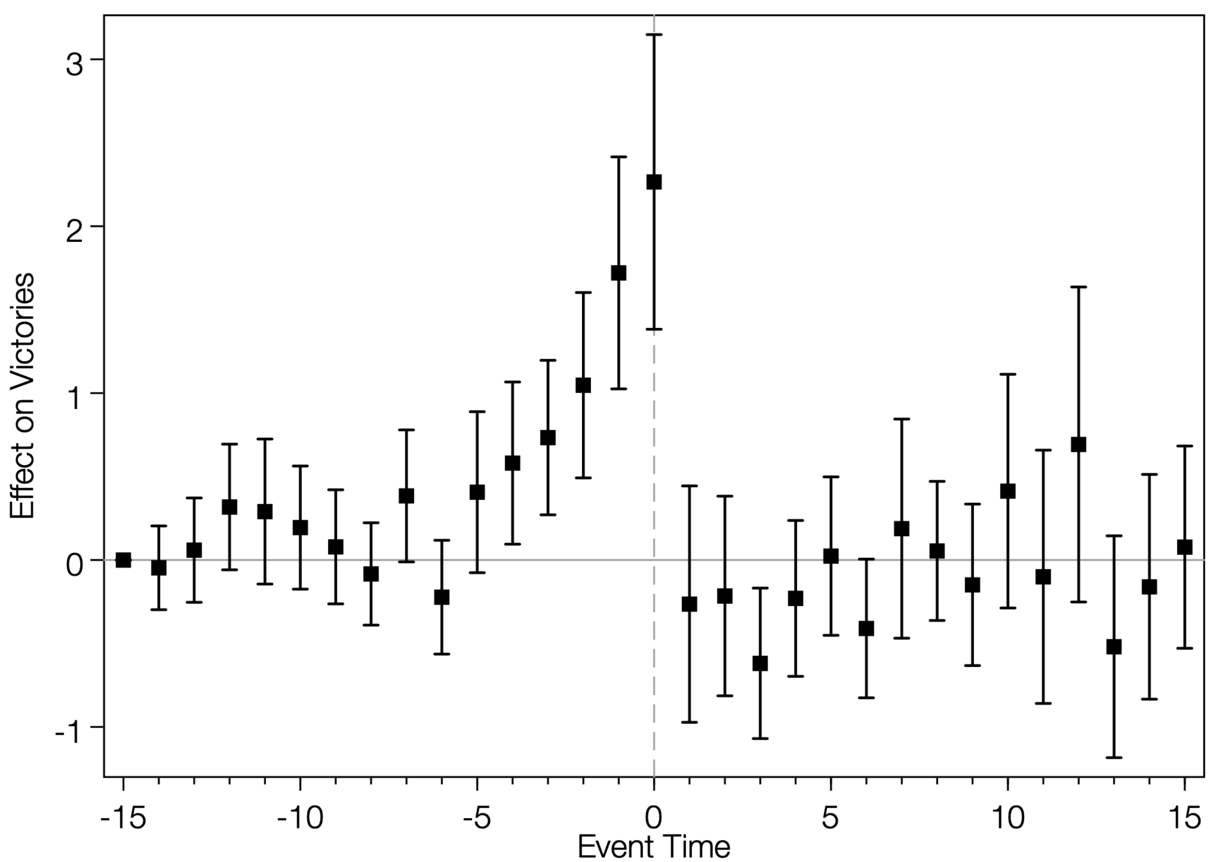
The event study in Figure 5 plots the μ_j terms and 95% confidence intervals from the mechanical-correction procedure. We interpret these terms as effort effects since we residualize monthly victories with a rich set of observables. Standard errors are clustered at the level of fighter squadrons. The figure shows effort increases in the run-up to the indicative quota of victories for the medal. Some 4-5 months before pilots meet the quota, their performance accelerates. The effect is substantial in magnitude—around one standard deviation in monthly performance at its peak. The total induced effort is the area under the treatment effect curve, which we estimate as 6.35 (s.e. = 0.62). By way of comparison, the median

$\overline{\mu} = 2.0$.

[12] Note that this correction results in regressions with approximately $p \cdot N = 25 \cdot 82,348$ observations. We cluster standard errors at the squadron level.

pilot in our data scores just three victories in their entire career, so each medal has an effect equivalent to deploying two additional pilots for the duration of their service (with vastly added cost in equipment and gasoline). Note that this result is driven by pilots who are much higher 'quality' than the median pilot. Immediately upon receiving the medal, effort drops sharply. This is important because it shows that "hot streaks"—when combat conditions facilitate a sequence of high-performance months—are not driving our results. Pilots exercise calculated precision in their exertions in response to the medal quota. Interestingly, we also find that pilot effort sharply falls once they receive the actual medal (even if the pilot is still below the quota, see Figure A7).

Figure 5: Pilot effort spikes around the medal quota - Medal 1



Notes: Estimated average treatment effects on victories per month for ace pilots (top 20% by ability, using the method from Mas and Moretti (2009)) as a function of their time relative to meeting the medal 1 quota. The event study model (2) includes 15 lags and leads of treatment; all further periods are absorbed into the 15th lag (lead), which is the reference period. The treatment occurs in period 0, so period 1 is the first period when pilots are fully treated. Estimation via OLS with controls for front, medal, and one lag of victories. Adjustments are made to remove the mechanical bias discussed in Section 2.1. Fixed effects are pilot and period. Standard errors clustered at the level of fighter squadrons with 95% confidence intervals indicated by the capped spikes.

We interpret the quickening pace of victories just before the pilot reaches the informal quota as increasing effort—as a pilot enters "the zone", every extra victory carries with it the additional promise of a highly coveted medal. The fact that the out-performance of pilots is heavily concentrated in periods when it can earn them individual glory further suggests that patriotic or ideological factors are unlikely to be the main motivating force in our sample.

The return of pilots to their baseline rhythm of performance after achieving the quota—and hence the medal—can reflect two effects. First, the immediate incentive of the medal is withdrawn, so pilots reduce their effort. The next medal quota is likely very far away. Second, medalists were often deployed on speaking tours and other propaganda and recruitment-related duties upon receipt of the medal. We cannot systematically distinguish between periods of inactivity due to these kinds of reassignments and periods of effort withdrawal or slackening off; therefore, we do not want to overinterpret these post-treatment effects.

Robustness

We perform several robustness tests to validate the result and aid interpretation. These additional tests are collectively summarized in Figure A8 alongside the baseline specification described above.

First, we estimate a parsimonious version of (2) without including covariates for medal status or the first lag of the outcome, ruling out concerns of Nickell (1981) bias in our fixed effects estimation. The overall pattern is unchanged though there is some carryover of the positive effects into period $j = 1$.

Second, we show the result is robust to estimating the effect only on ace pilots, defined as those pilots in the top skill quintile under the Mas and Moretti (2009) measure (approximately 1,000 pilots). Interpretability of the coefficient as effort in this group of pilots is easier, relative to the full sample, because we know from Figure 6 (discussed in detail below) that aces can carefully calibrate their effort and risk-taking in a way that lower-skill pilots are not. The differences versus the full sample results are minor, given that ace pilots are disproportionately represented in the treatment group.

Third, we show the robustness of the result to a different interpretation of the quota level.

The alternative coding was produced by a different reader of the source text (Obermaier, 1966) and takes a more conservative interpretation of the number of quota level changes that occur (though this was not a prompt to the coder). The baseline and alternative interpretations correlate with $\rho = 0.7399$ at the front $\times$ month level, and show only minor differences in the event study.

Third, we address a concern about how the informality of the quota affects our timing assumptions. Since the quota does not perfectly predict the timing of the medal award, we might be concerned that we are overestimating the anticipation effects because, in exceptional cases, pilots who are just below the victory quota are awarded the medal early due to significant accomplishments. This would be akin to demonstrating that significant accomplishments proceed to receiving a discretionary award, which would be unsurprising. We show that this channel does not explain our effect by showing an alternative specification where the event indicator is contingent on both meeting the quota *and* holding the medal. This cleans the treatment effects in periods $j < 0$ of the exceptional awardees described above because the treatment reflects only the "standard" case of getting the medal when meeting the quota. Under this definition, we cannot reject statistical equality with the baseline specification except in period $j = 0$. The effect here is negative, since pilots have their achievement affirmed by the medal in this period, so can begin to "slack off" earlier.

Alternative estimation using instrumental variables

A distinct approach to estimating the treatment effect uses instrumental variables. We can exploit the fact that the quota level varies over time and is orthogonal to the error terms ε_j of *individual* pilots. The quota level was adjusted by the High Command to ensure the medal's exclusivity, given combat conditions. In general, quotas were raised over time as combat intensified. Changes were determined for entire theaters of battle and are thus independent of the performance of any one pilot in each period, conditional on fixed effects for the front and theater. The extensive program of quota level changes is shown in Figure A1. In Appendix D, we show how a static version of the in zone effect can be recovered by instrumenting the medal quota proximity of an individual pilot with the front-wide changes in the quota level. The coefficients from this approach in Table A5 are highly comparable in magnitude to the main (corrected) event study approach, though this static approach is less well suited to identifying dynamic policy effects than our baseline event studies.

Effects on risk taking and exit

We study the risks associated with medal chasing while in zone of the medal quota. Despite generally high baseline levels of risk, we find no differentially higher levels of pilot exit while in zone.

We analyze our panel using Cox proportional hazards models, with the failure event being our exit indicator variable. Exits represent pilot death, serious injury, MIA/POW status, or other severe incapacitation. Table 1 presents the estimated hazard ratios from several specifications. The coefficient of interest is on the in zone treatment, defined as above to indicate being within fifteen victories of the next medal quota. Across a range of stratum specifications for front, fighter squadron, and medal status, and including a control for pilot quality, we find a consistent null effect of the treatment. We can reject even relatively small effects on pilot exit while in zone. Figure A9 provides visual evidence from the raw data to complement the regression analysis. We see that exit rates do not seem related to quota distance, in stark contrast to the clear patterns we document for both medal risk and victories.

We conclude that, while being in zone increases performance, it does so without significant increases in exit risk. We note that pilots who reach the zone are already of exceptional ability, so they have the leeway to increase their effort without taking excessive risk. It is for this reason that the organization prioritizes the motivation of these aces in the medal system.[13]

[13] Note that this is in contrast to the result in Ager et al. (2022), who find marked increases in mortality risk for medium-high quality pilots whose former squadron peer was publicly recognized with a mention in dispatches.

Table 1: Estimates of the inzone effect on exit

	Exit (0/1)				
	(1)	(2)	(3)	(4)	(5)
Inzone	1.156	1.109	1.031	0.940	1.079
	(0.104)	(0.101)	(0.091)	(0.085)	(0.111)
Pilot quality		1.025	1.025*	1.017	0.992
		(0.016)	(0.015)	(0.015)	(0.014)
N	80,361	77,624	77,624	77,624	77,624
χ^2	2.63	4.55	3.24	1.58	0.76
Strata			Front	Front Squadron	Front Squadron Medals

Notes: Table presents hazard ratios estimates from Cox regressions. The failure event is pilot exit. Inzone indicates that the pilot is within 15 victories of the next medal quota but has not earned the medal. Strata are defined based on combat front, fighter squadron ID, and the number of medals held. The baseline hazard rate is allowed to vary flexibly across strata. Standard errors (in parentheses) are clustered at the level of fighter squadrons.

3 The Organization's Problem

Having documented the increase in effort around medal quotas, we now turn to the organization's problem—to elicit effort repeatedly through the use of symbolic awards, despite the risks this entailed.

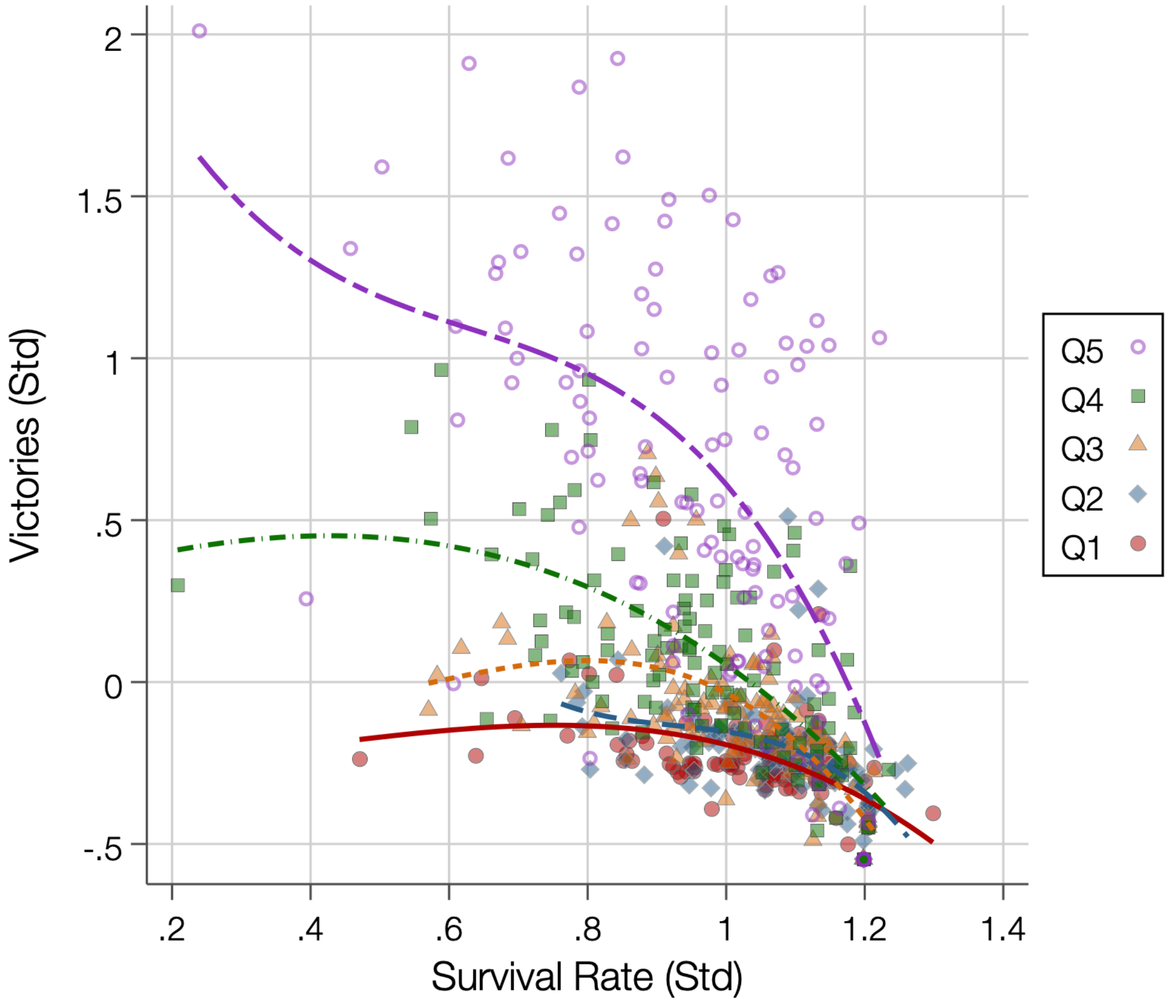
3.1 The cost of effort

Top pilots in the German air force contributed a large share of the overall victory tally—motivating flyers in this group was important. Of the total $54,835$ aerial victories in our data, 39.5% were scored by the 10% most able pilots by our Mas and Moretti (2009) measure of ability. Moreover, 8.6% were scored the top 1% of pilots alone.

At the same time, all pilots—including elite pilots—had to take substantial risks to win in aerial combat. Figure 6 illustrates the tradeoff as an empirical production possibilities frontier. We sort pilots into ability quintiles, and aggregate the data to the month × front × ability quintile level. On the x-axis, we plot the (normalized) average survival rate of pilots. On the y-axis, we plot the average victory rate. The further to the left an observation lies, the greater the risks they took; the higher an observation, the more enemy planes they destroyed.

The organization's problem is to incentivize pilots to perform as far to the top-right as possible in this space. Top pilots (lilac-dashed line) are best able to negotiate this trade-off: every unit of risk taken yields many more victories. Pilots in the second quintile are already less than half as productive as top pilots, destroying many fewer planes for every level of risk they took. For the third quintile, output per unit of risk again more than halves. Bottom-quintile pilots show a slight increase in victories as their risk rises. Even for top pilots, raising victory rates required greater risk. Creating new incentives for the highest-ability pilots to chase was thus essential to maintaining high aggregate performance.

Figure 6: Tradeoff between victory rate and risk by pilot quality quintile



Notes: Each observation is a month × front × ability quintile (Q5 contains the highest ability pilots). The figure shows averages of risk-taking (survival rate; one minus the exit rate) and performance (monthly victory rate). Outcomes are standardized at the year × front level such that $(1, 0)$ is the re-centered mean. Data are stratified by quintiles of the Mas and Moretti (2009) ability measure and summarized by cubic fits.

3.2 Medal cheapening

Typically, the first pilots receiving the medal (for their efforts) will be the best. As shown in Figure A10, the stock of medals already awarded grew continuously throughout the war.

At any moment, the signal value of the medal depends on the ability of the group of pilots who hold it. Initial awards may well have raised the salience and prestige of the medal, but this effect—as in the consumer good model of Pesendorfer—eventually flips sign: As time passes, progressively 'worse' pilots receive a medal, reducing the signal value of the award. This will leave those early, elite pilots dissatisfied with their (devalued) prize, potentially leading them to want something new.

To examine this question empirically, we rank the medalists using our econometrician's measure of ability, where rank #1 is the best pilot to receive the medal. As in the Pesendorfer model, this parameter is unobservable to both the peer pilots and the organization; pilots are motivated to obtain the medal precisely to signal their type. For each medal recipient today (the focal recipient), we study the rank of the next winner relative to today's winner. If the relative rank is decreasing in the order of medal receipts, then the medal is being awarded to progressively less outstanding pilots. We estimate the slope of the relationship between relative rank and relative medal order.

One limitation of our setting is that many medals (77%) were ceremonially awarded on non-unique days, so dates of the medal award cannot establish a precise order in which pilots earned the medal. To use all recipients in our data, we break ties using a permutation approach. We can generate an ordering by randomizing the order of the medals *within* the non-unique dates but retaining the medal order *across* dates. We call this the set of permuted orderings, which asymptotically contains the true order.[14].

$$Y_{i,w,p} = \theta + \beta O_{i,w,p} + \varepsilon_{i,w,p} \tag{3}$$

We fit the regression model in (3) and show both the linear and a non-parametric fit across $1,000$ permutations in Figure 7. For the recipient of the i^{th} medal, we create a dataset of the relative rank $Y_{i,w} = \text{rank}_i - \text{rank}_{i-w}$ of the pilots who received the medal in the window $[i-w, i+w]$, where w is the half-width of the window around the focal pilot which we set to 25. We then stack these datasets across all the medalists i and across all permutations p to estimate the coefficient on relative medal order O to be $\beta = -0.2444$. With standard errors

[14] Whether a true order exists is a conceptual point that we want to sidestep; our approach is robust to any method of resolving ties. If one does not resolve ties and lets the relative order of all pilots receiving medals on the same day be shared, the slope is $\beta = -0.2793$.

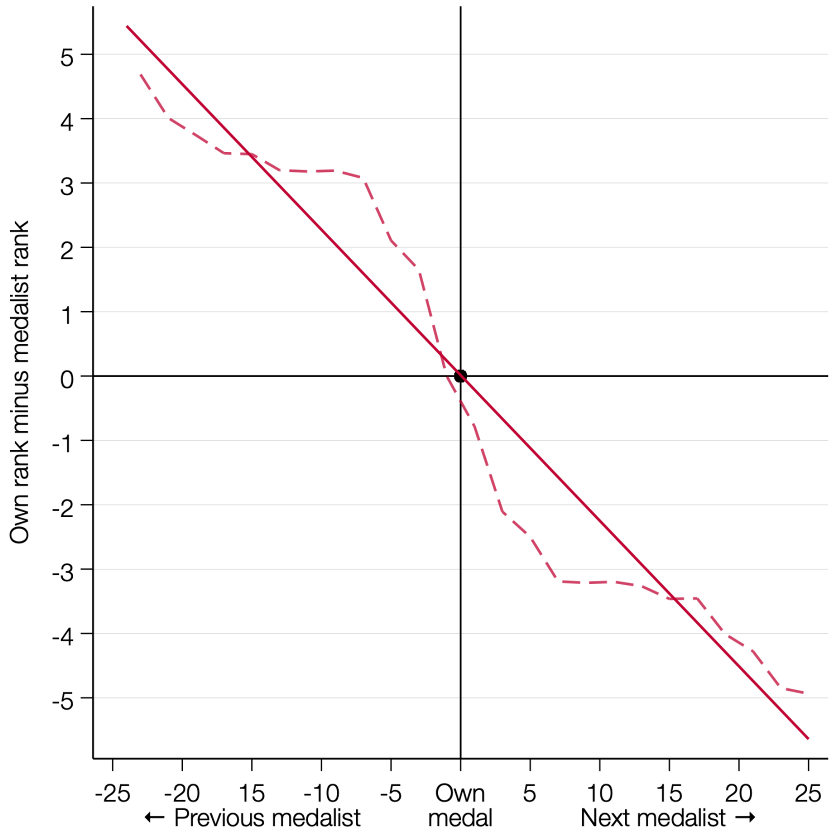
two-way clustered at the focal medalist and permutation level, we reject the one-sided null hypothesis of $\beta \geq 0$ with $p = 0.0009$. The medalist awarded medal 1 ten pilots after you is approximately two ranks lower-rated.

Our permutation approach also suggests a non-parametric mean of statistical inference; we compare the permuted orderings to a pure placebo set of orderings created by permuting the order of medalists across *all* dates, we can generate an entirely random ordering. We show the kernel density plots of the estimated slopes from the permuted orderings and the placebo orderings in Figure A11. The slopes from the permuted orderings are tightly clustered around the mean of the distribution, approximately the same mean recovered from estimating (3) on the stacked data. In contrast, the placebo orderings are loosely centered on zero. Our empirical $p = 0.0030$ is calculated as the share of the placebo ordering slopes more negative than the mean of the permuted ordering slopes. We show that the same separation of the placebo and permuted slopes, centered on a negative slope estimate, emerges from a range of window half-widths $w \in \{5, 10, ..., 50\}$ in Figure A12.

In a Pesendorfer-style model of trickle-down diffusion, high-status individuals are the first to adopt. In our setting, we need to rule out a mechanical reason for the "cheapening" of medals—the well of talent running dry over time, which would ensure that later awardees were less outstanding than their predecessors. Figure A13 plots the ability of new pilots (entrants in their first month as a pilot) in each month of the war at the 10th, 50th, and 90th percentiles. There are no time trends in entrant ability, which rules out the possibility that a change in talent composition is responsible for the observed cheapening effect.

We interpret our results from the perspective of an individual pilot. On average, those who received the medal before you were more highly skilled than you, making the medal a desirable target—you joined a highly exclusive club that will only just let you in. In contrast, those who followed you, while still elite, are lower-ranked pilots, diluting the prestige of your medal as time goes on. In this sense, medals share some characteristics with fashion trends, whose adoption initially raises demand before causing a decline in interest as "too many" of the lower-ranked types adopt a style. This pattern leads both pilots and fashionistas to chase the next status symbol.

Figure 7: Medal cheapening lowers the rank of later medal recipients

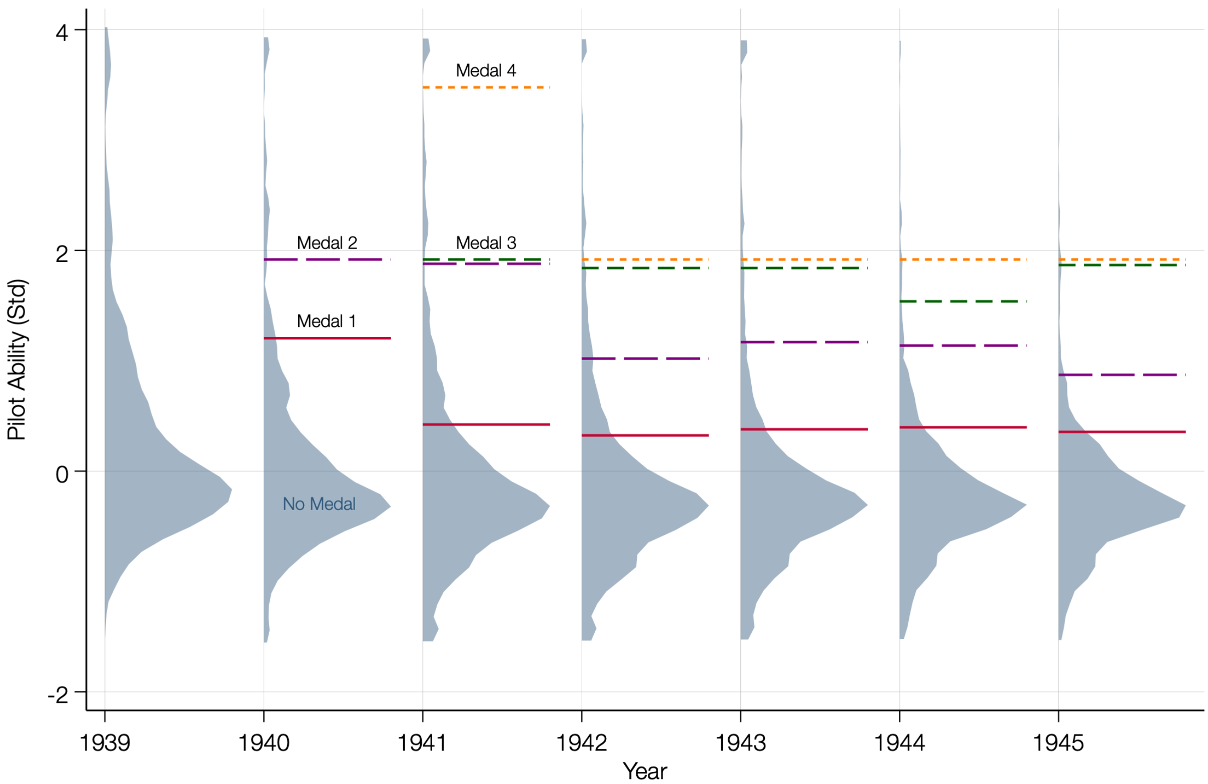


Notes: Figure shows the linear and local polynomial fit of relative rank and relative medal order for the 412 pilots who won medal 1, permuted $1,000$ times to resolve the ambiguous order of same-day medalists. Rank is in terms of ability among all pilots who are awarded the medal, where rank #1 is the best pilot. The x-axis shows the pilots who won the medal, ordered before and after the focal medalist at $x = 0$. The y-axis gives the rank of the focal pilot minus the rank of x^{th} medalist since the focal pilot was awarded the medal.

As a result of progressive medal cheapening, the quality of recipients declines over time. In Figure 8, we show the evolution of medal recipient quality relative to the overall distribution. For each year of the war, we plot the distribution of Mas-Moretti ability scores for pilots without any Knight's Cross medal, and indicate the median ability of medal holders in each category. The comparison of 1940 and 1941 shows sharp declines in the average quality of medal 1 holders. Between 1941 and 1942, the median quality of medal 2 holders exhibits a similar decline. When Medal 4 was introduced in 1941, the median quality of this group was markedly higher than that of Medal 3. However, just one year later, the quality of medal 4 recipients had declined markedly, with a value barely above that of medal 3 holders.

Overall, the German air force managed an elaborate system where quota changes and medal introductions counteracted the natural tendency of downward "quality drift", as in the Pesendorfer model. The separation of pilots of different abilities was broadly achieved, making medals useful signals, despite the tendency for quality of recipients to decline over time.

Figure 8: Quality of medal recipients declines over time

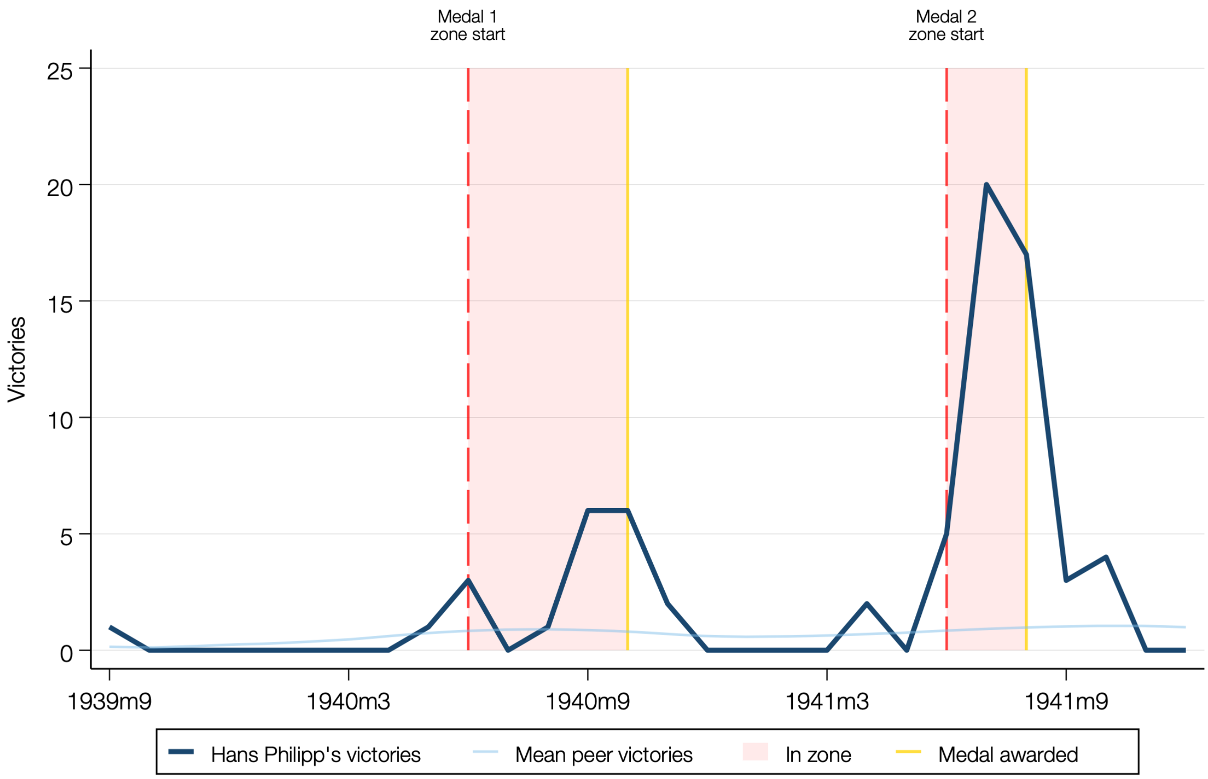


Notes: Figure shows the distribution of ability scores (Mas and Moretti, 2009) for each year for pilots who do not hold any medal. Horizontal lines indicate the median quality of medal recipients at each level of the Knight's Cross, from medal 1 through medal 4, once available.

4 Incentive Effects of the Medal 'Ladder'

Did the prospect of higher medals generate similar effort spikes as the initial award? Consider the case of Hans Phillip in Figure 9. Phillip is one of the top aces in our data, credited with over 200 career victories. The dark blue line shows his monthly victory tally, the light blue line, the average for his squadron peers. Dashed red lines indicate when the "zones" for medals 1 and 2 begin; gold lines indicate when the medals were awarded. We can see how Phillip's performance accelerates sharply - both in absolute terms and relative to his teammates - as he enters the range of cumulative victories that qualify him for medals 1 and 2. Immediately after each award, his performance slumps back toward the level of his peers.

Figure 9: Quotas and performance - Case study evidence



Notes: Figure shows the performance of fighter ace Hans Phillip (b. 1917) over time. The navy line gives the monthly victory score. The light blue line gives the mean score of squadron peers. The red vertical lines indicate when the pilot is in zone (within fifteen victories) of a medal quota. The gold lines indicate when a medal was awarded.

To go beyond anecdotal evidence, we use a combined event study design in Figure 10. We use the estimating equation in (4), which is a generalized form of (2) for two events; meeting the medals 1 and 2 quotas, respectively[15]. We adjust this estimation for the mechanical bias in the same way as described in Section 2. We denote the events using $m = 1, 2$. The model is fully saturated using 14 lags and leads for both treatments, plus binned pre- and post-treatment period dummies.[16] In addition, we include μ_{mid} to capture pilots between

[15]We plot the raw data for the medals 2 and 3 (in the fashion of Figure 2) in Figure A4.

[16]To achieve full saturation without introducing redundancy or overlapping effects between lags, we resolve periods in which

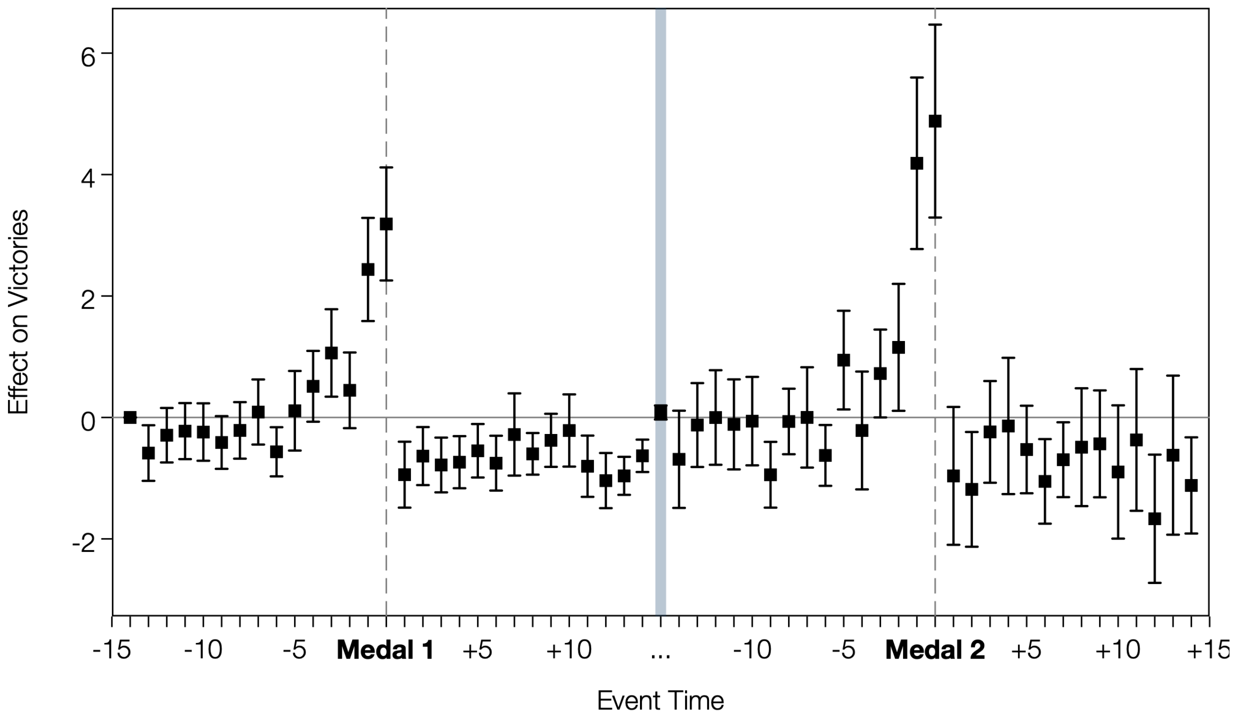
the two events but not captured by other lags/leads. The leave-out period is $j = -14$ of the first medal. The vector $\mathbf{X}_i$ is the same set of covariates as are included in (2).

$$\begin{aligned} Y_{i,t} = \sum_{m \in \{1,2\}} \sum_{j \in J} & \mu_{j,m} \mathbb{1}\{t - E_{i,m} = j\} \\ & + \mu_{mid} \mathbb{1}\{(t - E_{i,1} > j_{max}) \cdot (t - E_{i,2} < j_{min})\} \\ & + \beta \mathbf{X}_i + \varepsilon_{i,t} \end{aligned} \tag{4}$$

Pilot effort increases in the run-up to the second medal quota in a manner highly comparable—in duration and magnitude—to the first medal. The results suggest that pilots indeed reset their goals to the new target. We then estimate the static version of the event study design in Table A3, and extend the analysis to the third medal. Despite the limited number of awards, which causes our study to be underpowered, the third in zone treatment nonetheless exhibits the same pattern. Table A3 also presents the results of a pseudo-Poisson maximum likelihood estimation, given that our dependent variable is a count (Cameron and Trivedi, 2015). The results are highly consistent with and statistically stronger than our baseline OLS estimates.

the pilot satisfies two different lags by setting to 1 only the sequentially later lag.

Figure 10: Pilot performance spikes repeatedly at each medal quota



Notes: Estimated average treatment effects on victories per month as a function of time relative to meeting the relevant quota for ace pilots. The event study model (4) is fully saturated with 14 lags and leads of both treatments, less the 14th lead of the medal 1 quota as a reference period. Further periods are absorbed into the pre- and post-treatment dummies. Pilots between the two medals are absorbed into an additional lag denoted by "...". Observations that belong to multiple lags (e.g, 6 months after medal 1 is also 7 months before medal 2) are included only in the later lag (e.g, only in the 7th medal 2 lead). Estimation via OLS with controls for front and one lag of victories. Fixed effects are pilot and period. Adjustments are made to remove the mechanical bias discussed in Section 2.1. Standard errors clustered at the level of fighter squadrons with 95% confidence intervals indicated by the capped spikes.

4.1 The effects of medal introductions

Pilots entering the "zone" of victories that qualify them for an award increase their effort—a pattern that repeats for each new medal on the ladder. Reaching the zone, however, is itself a result of significant efforts and abilities beyond those of most pilots. Who cares for this rat race? When do they begin their efforts? Medal ladder extensions provide discrete moments in time that might incentivize any medal-minded pilot to exert effort—even if they are some way away from qualifying for the new medal. This response would be in line with a model of salience, whereby the introduction of new awards by the Luftwaffe to highlights the status benefits of medals, relative to the risk (Bordalo et al., 2012). Alternatively, the introductions might be moments of demotivation—the top gets further and further away.

Analyzing the treatment effect of medal introductions on pilots not in immediate "danger" of receiving the medal themselves is challenging. New medals are introduced only a few times. For the vast majority of pilots, these medal introductions were irrelevant, as they had little to no chance of ever meeting the quota and receiving the award—they are, in effect, untreated. Some pilots, on the other hand, were so highly skilled that for them, the prospect of (eventually) receiving the medal was an important source of motivation.

Ex-post, we know who the high-performance pilots were: those who reacted markedly to the prospect of a new medal as they entered the 'zone'. Since we are interested in the *timing* of performance changes, we restrict our analysis to this potential treatment group. We can then compare how introducing new awards affected "potential compliers" performance, depending on how much they were given to 'medal chasing' previously. Anecdotal evidence suggests that some soldiers were more keen on medals than others, jokingly being referred to as suffering from 'throat ache'—Knight's Crosses were worn around the neck (Hartmann, 2012). To account for additional, unobserved differences between the medal-sensitive group and other pilots, we use a latent variable approach in the spirit of synthetic control (Abadie and Vives-i Bastida, 2022; Liu et al., 2024).

The first step is to derive a measure of medal sensitivity. We use the causal tree method of Athey and Imbens (2016) as implemented by Yadlowsky et al. (2023) to calculate conditional average treatment effects (CATE). We use noble birth and officer status as heterogeneity-inducing covariates. Figure A14 shows the targeting operator characteristic (TOC), a measure of treatment effect heterogeneity. We find that these covariates predict substantial heterogeneity, with the most medal-sensitive pilots having average treatment effects 0.15 victories per month larger than the unconditional average treatment effect. Next, we use CATEs to examine the differential impact of medal introductions on victory rates by previous medal sensitivity. We call pilots in the top quintile of the in zone CATE distribution 'sensitive'.

To estimate the effect of medal introductions on sensitive pilots, we use the generalized synthetic control method of Xu (2017). First, we estimate an interacted fixed effects model using only the control group to recover $r = 1$[17] unobserved time-varying factors. Second, we use the pre-treatment outcomes for the treated group to infer unit-specific factor loadings for treated units. Third, we use the factors and factor loadings to predict counterfactual potential outcomes for the treated units in the post-treatment period. The model also includes standard two-way fixed effects and an indicator variable for the combat front. Intuitively, for each "medal-sensitive" pilot, we weigh appropriate control-group pilots with similar performance trajectories before the introduction of the medal into synthetic controls to make an out-of-sample prediction of the counterfactual outcome. Finally, we subtract the imputed counterfactual from each treated pilot's actual outcomes and compute the average treatment effect on the treated (ATT).

An advantage of this approach over standard two-way fixed effects is its ability to flexibly handle unobserved heterogeneity across units using latent factors, under the strong assumption that these factors apply equally to control and treatment units. This method is well-suited to our data since it can exploit the length of the pre-treatment period to estimate the factor loadings precisely.[18] Making within-period comparisons also sidesteps concerns about the endogeneity of medal introductions, whose precise timing might be in response to aggregate combat conditions (which would apply equally to medal-minded and other pilots). This estimator does not require a leave-out period, so ATTs will reflect the *absolute* difference in performance between medal-minded and control pilots. To facilitate a *relative* interpretation comparable to our other results, we re-center our estimates around a reference period.

We aim to recover pilots' responses to the possibility of a new medal. "Medal-minded" pilots are overwhelmingly *not* recipients of the higher medals whose introduction we study. Fewer than one in four recipients of medal 1 go on to win medal 2, and only one in eight win medal 3. We exclude, in each case, the pilots who later go on to win the medal being introduced. We also exclude pilots who are "topped out" and possess the previously top available medal to prevent measuring a cheapening effect on the old medal by the new (we focus on this mechanism in the next section). This allows us to interpret the effects as driven by the pull of the new award for pilots who are still some distance away from qualifying directly for the new awards. With these restrictions, our final analysis studies the behavior of 182 medal-minded pilots.

[17]Chosen under cross-validation to minimize mean squared percentage error in the pre-treatment period.

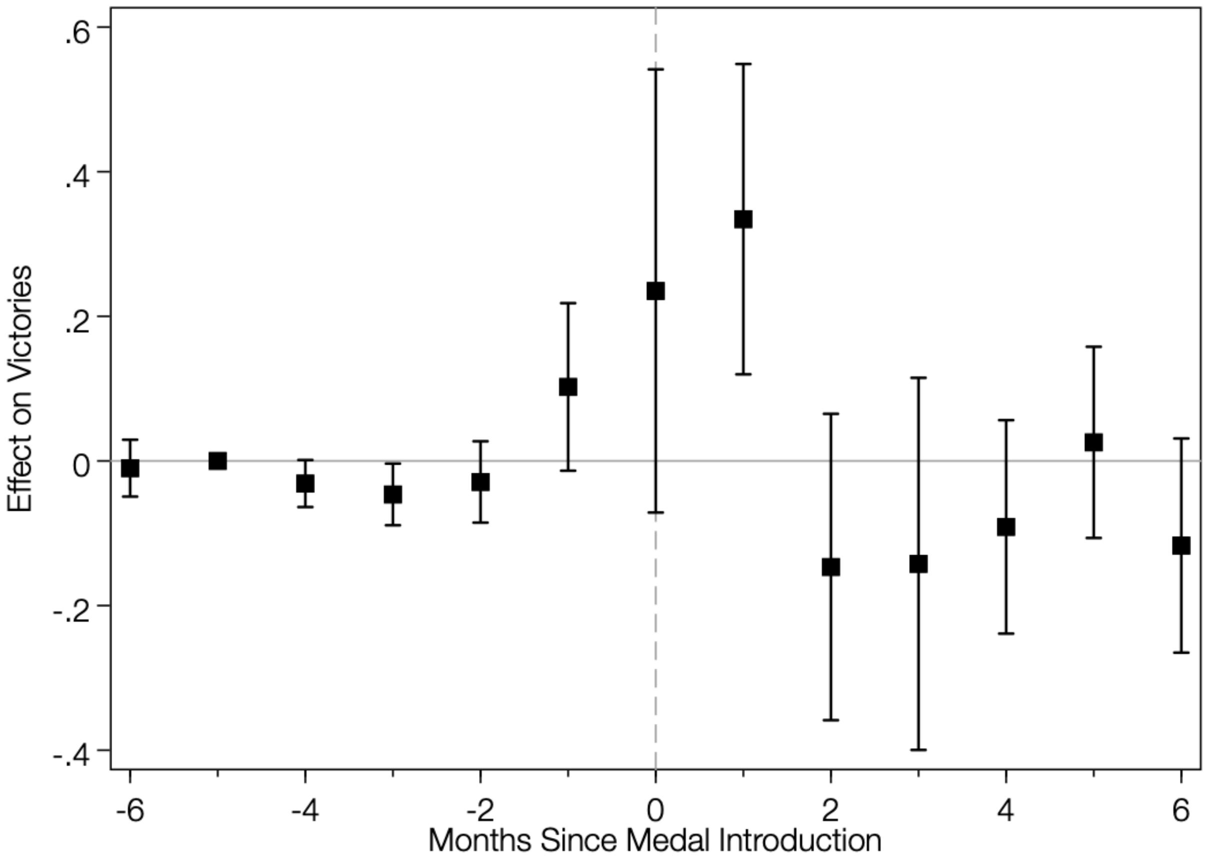
[18]Pilots with fewer than five observations are dropped from this analysis.

Figure 11 shows the results for the introductions of medal 2 in June 1940 and medals 3 and 4 (simultaneously) in July 1941. We focus on a short 6-month window centered at $t = 0$, since the events are only 13 months apart. We see evidence that medal-minded pilots are, at least initially, motivated by the announcement of a new, higher goal (even if it is far beyond their current realistic set of aspirations). The effect declines quickly over time and is zero after two months. This behavior is consistent with our earlier evidence on the quota effects—pilots increase effort around salient events, rather than smoothly increasing their level of effort in all periods. The effect size, however, is noticeably smaller than in the medal quota exercise. Pilots react to the news of the ladder extension by earning approximately 0.3 extra victory per month, for two months, before the effects fade. The small, positive treatment effect at $t = -1$ may reflect rumors or anticipation of the announcement, given that we take the introduction date as the day that the new administrative procedure was promulgated. This date may lag the moment at which officers and other pilots became aware of the impending ladder extension.

A standard robustness exercise with this approach involves selecting several pre-treatment periods as placebo periods to exclude from the model-fitting step. The estimated ATTs in these periods should be statistically insignificant if the model is correctly specified, alleviating concerns of over-fitting to the pre-period. We perform this exercise in Figure A15 and confirm that the estimator performs well.

We interpret these results as evidence supporting an organization's ability to repeatedly extract effort from its workers through a hierarchical system of incentives. Pilots respond to the creation of higher-order incentives, even when these targets are remote. Moreover, this suggests that pilots are aware of the entire system of medals, not just those immediately relevant to their career. However, a limitation of this exercise is the lack of true cross-sectional variation in exposure to the introductions, which necessitates making within-period comparisons against pilots who we assume will respond less to new medals.

Figure 11: Pilot effort increases when new medals are introduced



Notes: Figure shows the average treatment effect on the treated (ATT) of medal introductions on monthly victories for medal-sensitive pilots who are in the top quintile of conditional average treatment effects for the in zone treatment, using the remaining pilots as control units. Estimation via generalized synthetic control (Xu, 2017) with covariate matching on combat front and $r = 1$ latent factors. ATTs are re-centered relative to a reference period, $t = -5$. Bootstrap 95% percentile confidence intervals are indicated by the capped spikes.

5 Mechanisms and Magnitudes

In this section, we first examine whether there is additional, indirect evidence in favor of our proposed mechanism – status competition. We also discuss the aggregate effect, deriving a lower bound estimate of the total number of additional Allied planes destroyed because of the motivating effect of the German medal system

Heterogeneity of in zone response by pilot prestige

To understand what motivates pilots to exert extra effort when chasing medals, we examine alternative sources of prestige. Many (but far from all) medal-eligible pilots are officers and have the honor of command; some are noblemen and many may hold other, lower-tier medals. We hypothesize that the prospect of a highly prestigious medal may be more motivating for pilots who have few or no other sources of social standing. This would be in line with the prediction of status signaling models (see, e.g. Spence, 1973)—pilots with strong existing signals have weaker incentives to invest in costly additional signals.

We measure a pilot's prestige with the first principal component of dummies for holding (i) the German Cross (*Deutsches Kreuz*), (ii) a squadron commander's rank of second lieutenant (*Oberleutnant*) or higher, and (iii) a noble title.[19] Examples of high-status pilots according to this measure include Diethelm von Eichel-Streiber, a landed politician's son with a decorated service history in the Spanish Civil War, and Hubertus von Bonin, the son-in-law of a Prussian General and experienced squadron commander.

Using this treatment definition, we study the interaction between being "in zone" of the medal quota (as defined in the previous section) and having alternative sources of prestige. The first column of Table 2 replicates the static equivalent of the dynamic result in Figure 5—pilots work harder when close to the medal quota. Column 2 shows that having alternative sources of prestige is associated with higher performance. However, the interaction in Column 3 between prestige and in zone is strongly negative; high-status pilots appear less motivated by the prospect of the next medal than pilots without prestige markers. This exercise demonstrates that medals are *substitutes* rather than *complements* for other status signals. Table A2 shows that the result is robust to using various subsets of the prestige variables in the principal component construction. In line with this, Table A4 studies this interaction effect for medals 1–3 separately. These result echo Charles et al. (2009) in showing that acquiring visible signals of status is more important for individuals from lower-status socioeconomic groups.

[19]Noble status is time-invariant and indicated by the presence of the nobiliary particle *von, vom, zu, zum, von und zu, etc.* in the pilot's name. An *Oberleutnant* is the median and modal rank of the highest-ranked pilot in a fighter squadron in our data. In contrast with Western air forces, the German air force had a high proportion of NCOs flying fighters. Holding the German Cross was *not* a prerequisite for winning the Knight's Cross.

Table 2: High-status pilots show a smaller in zone response

	Victories		
	(1)	(2)	(3)
Inzone=1	0.438***	0.425***	0.483***
	(0.047)	(0.047)	(0.054)
Prestige (PCA)		0.221***	0.202***
		(0.039)	(0.038)
Inzone=1 × Prestige (PCA)			-0.088***
			(0.029)
Estimator	OLS	OLS	OLS
N	1,810,588	1,810,588	1,810,588
Inzone Pilots	573	573	573
FEs	Y	Y	Y
Outcome Mean	0.61	0.61	0.61
Outcome S.D	2.10	2.10	2.10

Notes: Table presents regression estimates. The outcome is individual monthly victories. In zone indicates that the pilot is within 15 victories of the medal quota. Prestige is the first principal component of dummies for holding a squadron commander's rank, the DK medal, and being from the nobility. Fixed effects for the pilot, combat front, and period are included in all models. Fixed effects for the pilot, combat front, and period are included in all models, as well as control variables for medal 1 and lagged victories. Adjustments are made to remove the mechanical bias discussed in Section 2.1. Standard errors clustered at the level of fighter squadrons.

5.1 Magnitudes

How much did the 'rat race' contribute to the overall performance of the German air force? Over the course of the war, German pilots made 54,835 victory claims. Already in the raw data, we can observe that pilots who got close to a quota spent approximately 20% of their time 'in zone', but scored 40% of their victories there (Figure A16).

To arrive at an aggregate figure, we can examine the number of victory claims by pilots passing through the 'zone' for the next medal, and sum up the area under the curve of victories as pilots approach the quota. In the raw data (as in Figure 2 and Figure A4), this yields an estimate of 11.84 victories for medal 1 and 17.67 for medal 2. After correcting for mechanical bias, using our estimator from Section 2, we compute the areas to be 6.35 (s.e. = 0.62) per in zone medal 1 pilot. Thus, 54% of the victories bump we document in the raw data can be causally attributed to the effect of the first medal quota. For medal 2, we compute the adjusted areas to be 7.32 (s.e. = 1.00), representing 41% of the raw total.

With these causal aggregates, we compute the overall magnitude by scaling with the number of treated pilots. In Table 3, we offer three approaches that yield comparable final results. These approaches differ in how many pilots we consider treated. In an optimistic view, all pilots who hit the zone are fully treated. This is likely an overestimate, since pilots who take risks exit at higher rates. In a conservative approach, only pilots who receive the medal they are in zone for are considered treated. For the same reason, this likely results in an underestimate, since pilots highly motivated by the zone may die before receiving the medal. Lastly, we consider a granular approach that weights each event-study coefficient μ_j in (2) by the number of pilots observed at that lag j[20]. In total, we find that the quota system caused $3,000-5,000$ extra victories, representing $5-10\%$ of total victories in our dataset. Compare this gain to the median pilot, who scored just 3 victories in their career. This implies that the immediate effect of the medal system — sharper incentives as the medal loomed large — added the same number of aerial victories as deploying around $1,000-1,600$ additional (average) pilots.

Finally, we note that these figures are plausibly a lower bound of the true total effect. By focusing only on the aspects of the medal system we can causally identify, we abstract from any impact on pilot behavior outside the zone. Since we find that these may also be significant (e.g, around introduction events in Section 4.1), the overall share of victories contributed by the medal system may have been even larger.

[20] In practice, these weights come very close to the number of pilots who meet the quota, which is a requirement for j to be defined.

Table 3: Aggregate magnitude of in zone effects

	Medal 1		Medal 2		Total Inzone Victories	
Description	AUC	Inzoners	AUC	Inzoners	Count	Share of Grand Total
Optimistic	6.35	623	7.32	222	5,581	10.18% (0.81%)
Conservative	6.35	418	7.32	104	3,416	6.23% (0.51%)
Granular	1814.94		1166.71		2,982	5.44% (0.44%)

Notes: Table presents calculations of the aggregate effect of the medal system, based on the treatment effect estimates from Section 2. In each row, we take an estimate of the area under the curve ("AUC") in the event study for each medal, and scale it by the number of treated pilots ("Inzoners"). We then compute the total count of victories induced by medals 1 and 2, and the share this count represents in the total number of victories in our dataset ($54,835$). Optimistic means we scale by the count of pilots who *ever* enter the zone for medal, regardless of completion. Conservative means we only scale by the number of pilots who win each medal. Granular means we weight each event-study coefficient μ_j from (2) by the number of pilots observed in the lag j. Standard errors in parentheses computed via the delta method, assuming independence of the AUC for each medal.

6 Conclusion

Can status concerns lead to a 'rat race' – and can organizations engineer such a competition for non-financial rewards to extract effort repeatedly? Despite the zero-sum nature of status competition, economic theory has predicted that such a system can be created if status concerns are sufficiently strong and ability differs substantially (Moldovanu et al., 2007). In this paper, we examine this question empirically in a dynamic setting. We provide direct evidence from WWII that one organization, the German air force, successfully created a ladder of highly coveted medals, providing status-based incentives for fighter pilots. The number of rungs on the ladder increased periodically as more workers earned the top awards. The successive medals potentially provided pilots with an incentive to keep exerting effort. Our study demonstrates that this system successfully increased performance.

We first employ a modified event-study design and data on the monthly performance of over 5,000 German pilots to demonstrate that pilot performance increases markedly as they approach the eligibility threshold for a medal. Success rates in aerial combat immediately slump after receiving the award, suggesting a sharp change in incentives and effort. We show that this incentive effect is weaker for pilots who already possess other markers of high social status, in line with predictions from a standard signaling game.

As pilots approach the successive medal in the ladder, the same effort path repeats itself. Pilots are not satisfied with achieving the already highly prestigious first medal. We employ a synthetic control design to demonstrate that status-conscious pilots—those who reacted the most to the prospect of winning a medal—also exert additional effort when new medals are introduced during the war.

We are guided by the predictions of fashion cycle models (Pesendorfer, 1995). While the organization in our study is creating artificial scarcity, not selling novelty, the underlying mechanism is related: The best pilots win the medal first, while low types follow. This cheapening provides a rationale for new, rarer variants, which in turn motivate the high-type pilots to participate in the rat race, chasing the next (new) medal.

While extensions of the medal ladder could theoretically be demotivating, our empirical results show that a hierarchical system of status-based incentives can repeatedly elicit effort due to status concerns. The dynamics of the scheme generate a rat race among workers. The net effect of the medal ladder was substantial—we estimate that it added between 3,000 and 5,000 victories to German pilots' tally during the war (equivalent to 5-10% of the total). The organization's credibility in our study—the national armed forces during wartime— gives it leeway to construct status symbols that are ex ante desirable. Not every organization will enjoy this capability, resulting in varying degrees of rat races that should be explored in other settings.

References

Abadie, Alberto and Jaume Vives i Bastida, "Synthetic Controls in Action," March 2022. arXiv:2203.06279 [econ, stat].

Ager, Philipp, Leonardo Bursztyn, Lukas Leucht, and Hans-Joachim Voth, "Killer Incentives: Rivalry, Performance and Risk-Taking among German Fighter Pilots, 1939–45," *The Review of Economic Studies*, October 2022, *89* (5), 2257–2292.

Akerlof, George, "The Economics of Caste and of the Rat Race and Other Woeful Tales," *The Quarterly Journal of Economics*, 1976, *90* (4), 599–617. Publisher: Oxford University Press.

Akerlof, George A. and Rachel E. Kranton, "Identity and the Economics of Organizations," *Journal of Economic Perspectives*, March 2005, *19* (1), 9–32.

Allen, Eric J., Patricia M. Dechow, Devin G. Pope, and George Wu, "Reference-Dependent Preferences: Evidence from Marathon Runners," *Management Science*, April 2016. Publisher: INFORMS.

Ashenfelter, Orley, "Estimating the Effect of Training Programs on Earnings," *The Review of Economics and Statistics*, 1978, *60* (1), 47–57. Publisher: The MIT Press.

Ashraf, Nava and Oriana Bandiera, "Social Incentives in Organizations," *Annual Review of Economics*, 2018, *10* (1), 439–463. _eprint: https://doi.org/10.1146/annurev-economics-063016-104324.

_ , _ , and Scott S. Lee, "Awards Unbundled: Evidence from a Natural Field Experiment," *Journal of Economic Behavior & Organization*, April 2014, *100*, 44–63.

Athey, Susan and Guido Imbens, "Recursive partitioning for heterogeneous causal effects," *Proceedings of the National Academy of Sciences*, July 2016, *113* (27), 7353–7360.

Besley, Timothy and Maitreesh Ghatak, "Status Incentives," *American Economic Review*, April 2008, *98* (2), 206–211.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, "Salience Theory of Choice Under Risk," *The Quarterly Journal of Economics*, August 2012, *127* (3), 1243–1285.

Borusyak, Kirill and Peter Hull, "Nonrandom Exposure to Exogenous Shocks," *Econometrica*, 2023, *91* (6), 2155–2185.

Bottan, Nicolas Luis and Ricardo Perez Truglia, "Deconstructing the Hedonic Treadmill: Is Happiness Autoregressive?," *The Journal of Socio-Economics*, May 2011, *40* (3), 224–236.

Brickman, P. and D. Campbell, "Hedonic relativism and planning the good society," in "Adaptation-level theory : a Symposium [on Adaptation-Level (AL) Theory held at the University of Massachusetts at Amherst, on May 7-9, 1970]" 1971.

Bungay, Stephen, *The Most Dangerous Enemy: A History of the Battle of Britain*, Aurum, 2010.

Bursztyn, Leonardo, Bruno Ferman, Stefano Fiorin, Martin Kanz, and Gautam Rao, "Status Goods: Experimental Evidence from Platinum Credit Cards*," *The Quarterly Journal of Economics*, August 2018, *133* (3), 1561–1595.

Bénabou, Roland and Jean Tirole, "Identity, Morals, and Taboos: Beliefs as Assets *," *The Quarterly Journal of Economics*, May 2011, *126* (2), 805–855.

Caldwell, Donald, *JG 26 Luftwaffe Fighter Wing War Diary: 1943-45*, Stackpole Books, October 2012. Google-Books-ID: Zo8h7hNePhUC.

Calonico, Sebastian, Matias D. Cattaneo, and Rocío Titiunik, "Optimal Data-Driven Regression Discontinuity Plots," *Journal of the American Statistical Association*, October 2015, *110* (512), 1753–1769.

Cameron, A. Colin and Pravin K. Trivedi, "Count Panel Data," in Badi H. Baltagi, ed., *The Oxford Handbook of Panel Data*, Oxford University Press, January 2015, p. 0.

Charles, Kerwin Kofi, Erik Hurst, and Nikolai Roussanov, "Conspicuous Consumption and Race*," *The Quarterly Journal of Economics*, May 2009, *124* (2), 425–467.

Cullen, Zoë and Ricardo Perez-Truglia, "How Much Does Your Boss Make? The Effects of Salary Comparisons," *Journal of Political Economy*, March 2022, *130* (3), 766–822. Publisher: The University of Chicago Press.

Eisenkopf, Gerald and Sabrina Teyssier, "Envy and Loss Aversion in Tournaments," *Journal of Economic Psychology*, 2013, *34*, 240–255.

Frederick, Shane and George Loewenstein, "Hedonic adaptation," in "Well-being: The foundations of hedonic psychology," New York, NY, US: Russell Sage Foundation, 1999, pp. 302–329.

Frey, Bruno S., "Awards as Compensation," *European Management Review*, 2007, *4* (1), 6–14.

_ **and Jana Gallus**, "Towards an Economics of Awards," *Journal of Economic Surveys*, 2017, *31* (1), 190–200.

Frieser, Karl-Heinz, "Die Schlacht im Kursker Bogen," in "Das Deutsche Reich und der Zweite Weltkrieg," Vol. 8, Stuttgart: Deutsche Verlags-Anstalt, 2007, pp. 83–211.

Genakos, Christos and Mario Pagliero, "Interim Rank, Risk Taking, and Performance in Dynamic Tournaments," *Journal of Political Economy*, 2012, *120* (4), 782–813.

Genicot, Garance and Debraj Ray, "Aspirations and Economic Behavior," *Annual Review of Economics*, 2020, *12* (1), 715–746. _eprint: https://doi.org/10.1146/annurev-economics-080217-053245.

Gill, David and Rebecca Stone, "Fairness and Desert in Tournaments," *Games and Economic Behavior*, 2010, *69* (2), 346–364.

Hamilton, Stephen F., Timothy J. Richards, Aric P. Shafran, and Kathryn N. Vasilaky, "Farm labor productivity and the impact of mechanization," *American Journal of Agricultural Economics*, 2022, *104* (4), 1435–1459.

Hartmann, Christian, *Wehrmacht im Ostkrieg: Front und militärisches Hinterland 1941/42*, Vol. 75, Walter de Gruyter, 2012.

Heath, Chip, Richard P. Larrick, and George Wu, "Goals as Reference Points," *Cognitive Psychology*, February 1999, *38* (1), 79–109.

Hull, C. L., "The Rat's Speed-of-locomotion Gradient in the Approach to Food," *Journal of Comparative Psychology*, 1934, *17* (3), 393–422.

Imas, Alex and Kristof Madarasz, "Superiority-Seeking and the Preference for Exclusion," August 2022.

Kamenica, Emir, “Behavioral Economics and Psychology of Incentives,” *Annual Review of Economics*, 2012, *4* (1), 427–452.

Kivetz, Ran, Oleg Urminsky, and Yuhuang Zheng, “The Goal-Gradient Hypothesis Resurrected: Purchase Acceleration, Illusionary Goal Progress, and Customer Retention,” *Journal of Marketing Research*, February 2006, *43* (1), 39–58.

Kosfeld, Michael and Susanne Neckermann, “Getting More Work for Nothing? Symbolic Awards and Worker Performance,” *American Economic Journal: Microeconomics*, August 2011, *3* (3), 86–99.

Landers, Renée M., James B. Rebitzer, and Lowell J. Taylor, “Rat Race Redux: Adverse Selection in the Determination of Work Hours in Law Firms,” *The American Economic Review*, 1996, *86* (3), 329–348. Publisher: American Economic Association.

Lazear, Edward P., “Compensation and Incentives in the Workplace,” *Journal of Economic Perspectives*, August 2018, *32* (3), 195–214.

_ **and Sherwin Rosen**, “Rank-Order Tournaments as Optimum Labor Contracts,” *Journal of Political Economy*, 1981, *89* (5), 841–864. Publisher: University of Chicago Press.

Liu, Licheng, Ye Wang, and Yiqing Xu, “A Practical Guide to Counterfactual Estimators for Causal Inference with Time-Series Cross-Sectional Data,” *American Journal of Political Science*, 2024, *68* (1), 160–176. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ajps.12723.

Mas, Alexandre and Enrico Moretti, “Peers at Work,” *American Economic Review*, March 2009, *99* (1), 112–145.

Misra, Sanjog and Harikesh S. Nair, “A Structural Model of Sales-force Compensation Dynamics: Estimation and Field Implementation,” *Quantitative Marketing and Economics*, September 2011, *9* (3), 211–257.

Moldovanu, Benny, Aner Sela, and Xianwen Shi, “Contests for Status,” *Journal of Political Economy*, April 2007, *115* (2), 338–363. Publisher: The University of Chicago Press.

Murray, Williamson, *Strategy For Defeat: The Luftwaffe, 1933-1945*, Pickle Partners Publishing, 2015.

Nickell, Stephen, "Biases in Dynamic Models with Fixed Effects," *Econometrica*, 1981, *49* (6), 1417–1426.

Obermaier, Ernst, *Die Ritterkreuzträger der Luftwaffe: Jagdflieger 1939-1945*, Mainz: Verlag Dieter Hoffman, 1966. Google-Books-ID: 6yfQzwEACAAJ.

O'Donoghue, Ted and Matthew Rabin, "Present Bias: Lessons Learned and to Be Learned," *American Economic Review*, May 2015, *105* (5), 273–279.

Perez-Truglia, Ricardo, "On the Causes and Consequences of Hedonic Adaptation," *Journal of Economic Psychology*, December 2012, *33* (6), 1182–1192.

_ , "The Effects of Income Transparency on Well-Being: Evidence from a Natural Experiment," *American Economic Review*, April 2020, *110* (4), 1019–1054.

Pesendorfer, Wolfgang, "Design Innovation and Fashion Cycles," *The American Economic Review*, 1995, *85* (4), 771–792. Publisher: American Economic Association.

Prendergast, Canice, "The Provision of Incentives in Firms," *Journal of Economic Literature*, March 1999, *37* (1), 7–63.

Ray, Debraj and Arthur Robson, "Status, Intertemporal Choice, and Risk-Taking," *Econometrica*, 2012, *80* (4), 1505–1531.

Rayo, Luis, "Monopolistic Signal Provision†," *The B.E. Journal of Theoretical Economics*, January 2013, *13* (1), 27–58. Publisher: De Gruyter.

Simkin, Mikhail V and Vwani P Roychowdhury, "Theory of Aces: Fame by Chance or Merit?," *Journal of Mathematical Sociology*, 2006, *30* (1), 33–42. Publisher: Taylor & Francis.

Spence, Michael, "Job Market Signaling," *The Quarterly Journal of Economics*, 1973, *87* (3), 355–374. Publisher: Oxford University Press.

Williamson, Gordon and Ramiro Bujeiro, *Knight's Cross and Oak-Leaves Recipients 1939–40* Elite, 1st ed., London, United Kingdom: Bloomsbury Publishing, July 2004.

Xu, Yiqing, "Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models," *Political Analysis*, January 2017, *25* (1), 57–76. Publisher: Cambridge University Press.

Yadlowsky, Steve, Scott Fleming, Nigam Shah, Emma Brunskill, and Stefan Wager, "Evaluating Treatment Prioritization Rules via Rank-Weighted Average Treatment Effects," 2023.

Appendix for

Never Enough: Dynamic Status Incentives in Organizations

Leonardo Bursztyn
Ewan Rawcliffe
Hans-Joachim Voth

For Online Publication

List of Tables

A1 Summary statistics for key analysis variables A-2
A2 Robustness of in zone response heterogeneity to prestige measures A-3
A3 Additional estimates of the in zone effect by medal A-4
A4 Heterogeneity of the in zone response for higher medals A-5
A5 Using quota changes as an instrument for the in zone response A-6
A6 Balance of pilot attributes based on inclusion in ability estimation A-27

List of Figures

A1 Medal quotas are adjusted throughout the war A-7
A2 The Knight's Cross medal A-8
A3 Process of medal award A-9
A4 Raw victory rates over higher medal event time A-10
A5 Trends in other variables over medal event time A-11
A6 Comparison of naive and adjusted estimators in synthetic data A-12
A7 Victory rates around quota A-13
A8 Pilot effort spikes around the medal quotas, robustness exercises A-14
A9 Pilot exit risk does not track quota distance A-15
A10 The stock of medals awarded rises throughout the war A-16
A11 The estimated rate of cheapening is robust to permutations in medal order . A-17
A12 The estimated rate of cheapening is robust to the width of the comparison . A-18
A13 Ability of entrant pilots over time is constant A-19
A14 In zone treatment effects vary substantially across pilots A-20
A15 Pilot effort increases when new medals are introduced: placebo A-21
A16 Time 'in zone' contributes disproportionately to overall victory tally A-22
A17 Pilot ability estimation: squadron composition variability A-24
A18 Pilot ability estimation: peer combination frequency A-25
A19 Pilot ability estimation: peer fixed effects and model fit A-26
A20 Local polynomial estimates of zone entry A-30
A21 Quota changes and performance: Case study evidence A-31

A Additional Figures and Tables

Table A1: Summary statistics for key analysis variables

	N	mean	s.d	min	p50	p90	max
A. Pilot-month level							
Victories	82,348	0.67	2.15	0	0	2.00	68
Exit (0/1)	82,348	0.05	0.22	0	0	0	1
Eastern Front (0/1)	82,348	0.36	0.48	0	0	1	1
Experience (months)	82,348	18.07	15.27	1	14.00	42.00	69
Holds Medal 1 (0/1)	82,348	0.09	0.28	0	0	0	1
Switched Squadron (0/1)	82,348	0.04	0.19	0	0	0	1
Commander (0/1)	76,110	0.25	0.43	0	0	1	1
Distance to Medal 1 Quota	82,348	-33.66	25.15	-100	-37.00	-13.00	303
B. Pilot level							
Total Victories	5,081	10.79	23.01	1	3.00	27.00	343
Survives War (0/1)	5,081	0.19	0.39	0	0	1	1
Entry Period	5,081	38.03	18.99	1	42.00	60	68
Pilot Ability	4,664	0.04	1.34	-18.79	-0.13	1.24	18.07
Total Experience (months)	5,081	16.21	17.52	1	9.00	46.00	69
Nobility (0/1)	5,081	0.02	0.13	0	0	0	1
Holds DK Medal (0/1)	5,081	0.16	0.36	0	0	1	1
Receives Medal 1 (0/1)	5,081	0.08	0.27	0	0	0	1
Receives Medal 2 (0/1)	5,081	0.02	0.14	0	0	0	1
Receives Medal 3 (0/1)	5,081	0.01	0.08	0	0	0	1

Notes: Table presents summary statistics for key analysis variables.

Table A2: Robustness of in zone response heterogeneity to prestige measures

	Baseline	DK+Commander	Noble+Commander	Noble+DK
	(1)	(2)	(3)	(4)
Inzone=1	0.483***	0.488***	0.446***	0.441***
	(0.054)	(0.056)	(0.049)	(0.048)
Prestige (PCA)	0.202***	0.199***	0.216***	
	(0.038)	(0.038)	(0.039)	
Inzone=1 ×	-0.088***	-0.091***	-0.065**	-0.035
Prestige (PCA)	(0.029)	(0.032)	(0.025)	(0.028)
Estimator	OLS	OLS	OLS	OLS
N	1,810,588	1,810,588	1,810,588	1,810,588
Inzone Pilots	573	573	573	573
FEs	Y	Y	Y	Y
Outcome Mean	0.61	0.61	0.61	0.61
Outcome S.D	2.10	2.10	2.10	2.10

Notes: Table presents regression estimates. The outcome is individual monthly victories. Inzone indicates that the pilot is within 15 victories of the first medal quota without the medal. Prestige is the first principal component of dummies for holding a squadron commander's rank or higher, the DK medal, and being from the nobility. Columns 2–4 sequentially drop one variable from the PCA. Fixed effects for the pilot, combat front, and period are included in all models, as well as control variables for medal 1 and lagged victories. The main effect of prestige in Column 4 is absorbed by the pilot fixed effect because the dummies for Noble and DK are time-invariant. Standard errors clustered at the level of fighter squadrons.

Table A3: Additional estimates of the in zone effect by medal

	Panel A: Victories			
	(1)	(2)	(3)	(4)
Inzone for medal 1	0.498***			0.507***
	(0.056)			(0.056)
Inzone for medal 2		0.291***		0.306***
		(0.062)		(0.062)
Inzone for medal 3			0.322	0.346*
			(0.202)	(0.203)
Estimator	OLS	OLS	OLS	OLS
N	1,954,238	1,954,238	1,954,238	1,954,238
Inzone Pilots	610	214	47	616
FEs	Y	Y	Y	Y
Outcome Mean	0.61	0.61	0.61	0.61
Outcome S.D	2.12	2.12	2.12	2.12

	Panel B: Victories			
	(1)	(2)	(3)	(4)
Inzone for medal 1	0.242***			0.266***
	(0.022)			(0.023)
Inzone for medal 2		0.149***		0.190***
		(0.024)		(0.025)
Inzone for medal 3			0.095*	0.150***
			(0.054)	(0.056)
Estimator	PPML	PPML	PPML	PPML
N	1,743,196	1,743,196	1,743,196	1,743,196
Inzone Pilots	610	214	47	616
FEs	Y	Y	Y	Y
Outcome Mean	0.69	0.69	0.69	0.69
Outcome S.D	2.23	2.23	2.23	2.23

Notes: Table presents regression estimates. The outcome is the monthly victory rate. Inzone # indicates that the pilot is within 15 victories of the medal quota #, but without the medal #. Panel A presents ordinary least squares estimates. Panel B estimates the same models using pseudo-Poisson maximum likelihood (PPML). Singleton observations and observations separated by a fixed effect are excluded in PPML. Fixed effects for pilot, combat front, and period are included in all models, as well as control variables for medal 1 and lagged victories. Standard errors clustered at the level of fighter squadrons.

Table A4: Heterogeneity of the in zone response for higher medals

	Medal 1			Medal 2			Medal 3		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Inzone=1	0.490***	0.479***	0.532***	0.256***	0.243***	0.256**	0.300	0.284	0.761
	(0.055)	(0.054)	(0.068)	(0.061)	(0.060)	(0.100)	(0.218)	(0.216)	(0.578)
Prestige (PCA)		0.223***	0.211***		0.226***	0.216***		0.228***	0.229***
		(0.039)	(0.039)		(0.039)	(0.038)		(0.039)	(0.039)
Inzone=1 ×			-0.094**			-0.023			-0.350
Prestige (PCA)			(0.037)			(0.065)			(0.316)
Estimator	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
N	1,810,588	1,810,588	1,810,588	1,810,588	1,810,588	1,810,588	1,810,588	1,810,588	1,810,588
Inzone Pilots	567	567	567	196	196	196	36	36	36
FEs	Y	Y	Y	Y	Y	Y	Y	Y	Y
Outcome Mean	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61
Outcome S.D	2.10	2.10	2.10	2.10	2.10	2.10	2.10	2.10	2.10

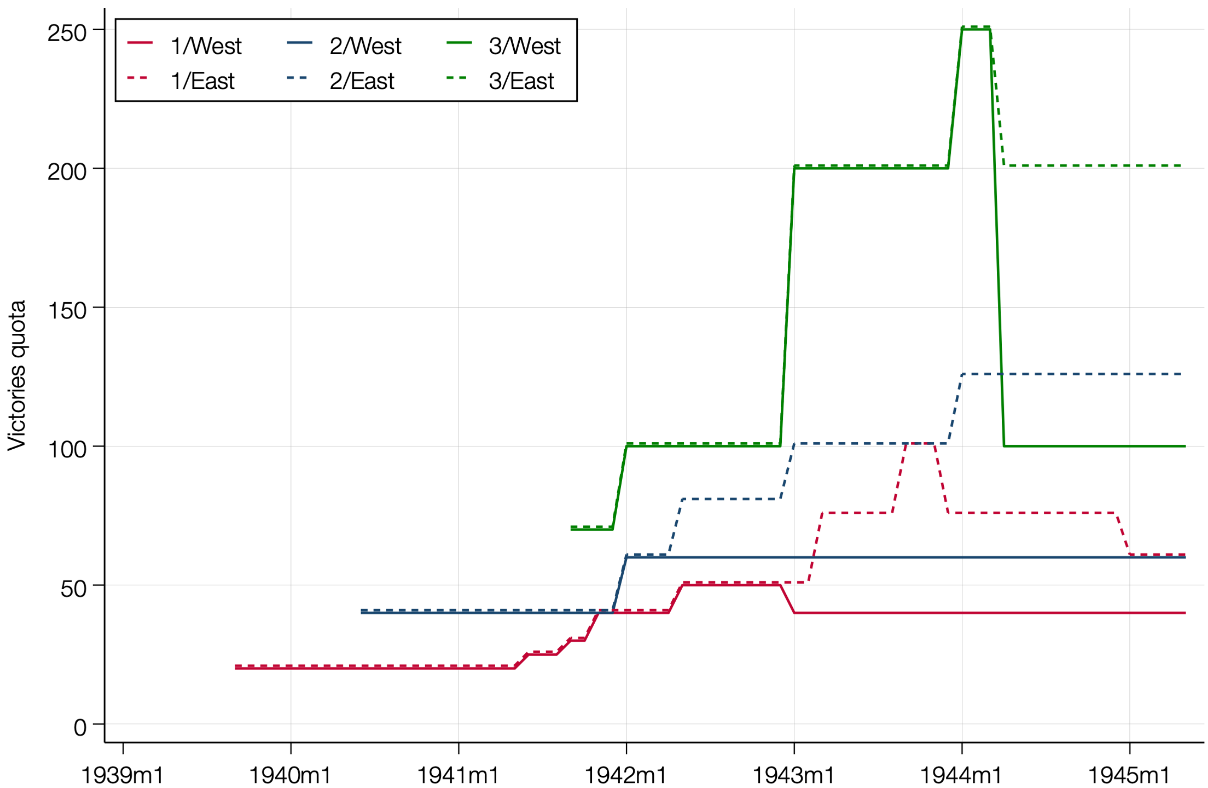
Notes: Table presents regression estimates. The outcome is individual monthly victories. Inzone indicates that the pilot is within 15 victories of the first medal quota without the medal. Prestige is the first principal component of dummies for holding an officer rank, the DK medal, and being from the nobility. Fixed effects for the pilot, combat front, and period are included in all models. Standard errors clustered at the level of fighter squadrons.

Table A5: Using quota changes as an instrument for the in zone response

	Victories					
	(1)	(2)	(3)	(4)	(5)	(6)
In Zone	1.477***	0.990***		2.750***	3.162**	4.544***
	(0.119)	(0.091)		(0.801)	(1.430)	(1.235)
Lag Victories		0.311***				0.261***
		(0.012)				(0.021)
Awarded KCR 0/1		-0.338***				-0.385***
		(0.098)				(0.145)
Quota Increase			-0.258***			
			(0.077)			
Estimator	OLS	OLS	OLS	2SLS	2SLS	2SLS
N	74,729	74,729	74,729	74,729	71,291	71,291
Inzone Pilots	616	616	616	616	430	430
FEs	Y	Y	Y	Y	Y	Y
Outcome Mean	0.61	0.61	0.61	0.61	0.53	0.53
Outcome S.D	2.12	2.12	2.12	2.12	1.88	1.88
KPW F-Stat				119.78	64.03	78.23
AR p-value				0.00	0.03	0.00
Std Coef (SE)	0.71 (0.04)	0.56 (0.03)	-0.08 (0.04)	0.81 (0.40)	0.87 (0.71)	1.33 (0.63)

Notes: Table presents regression estimates. The outcome is individual monthly victories. Inzone indicates ace pilots who are within 15 victories of their next medal quota. Lag Victories is the first lag of the outcome variable. Awarded KCR is a dummy capturing whether the pilot holds medal 1. Quota Increase is a front×period×medal dummy variable measuring whether the level of the medal quota this period is higher than last. The 2SLS estimates in columns 4–6 use Quota Increase as an instrument for Inzone. Columns 5 and 6 restrict the sample to exclude in zone pilots with in zone squadron peers. Std Coef reports the lead coefficient from each model where the outcome victory rate has been standardized at the front×period level. Fixed effects for the pilot, combat front, and period are included in all models. The Kleibergen-Paap rk Wald F statistic and Anderson-Rubin p-values are reported for 2SLS estimates. Standard errors clustered at the level of fighter squadrons.

Figure A1: Medal quotas are adjusted throughout the war

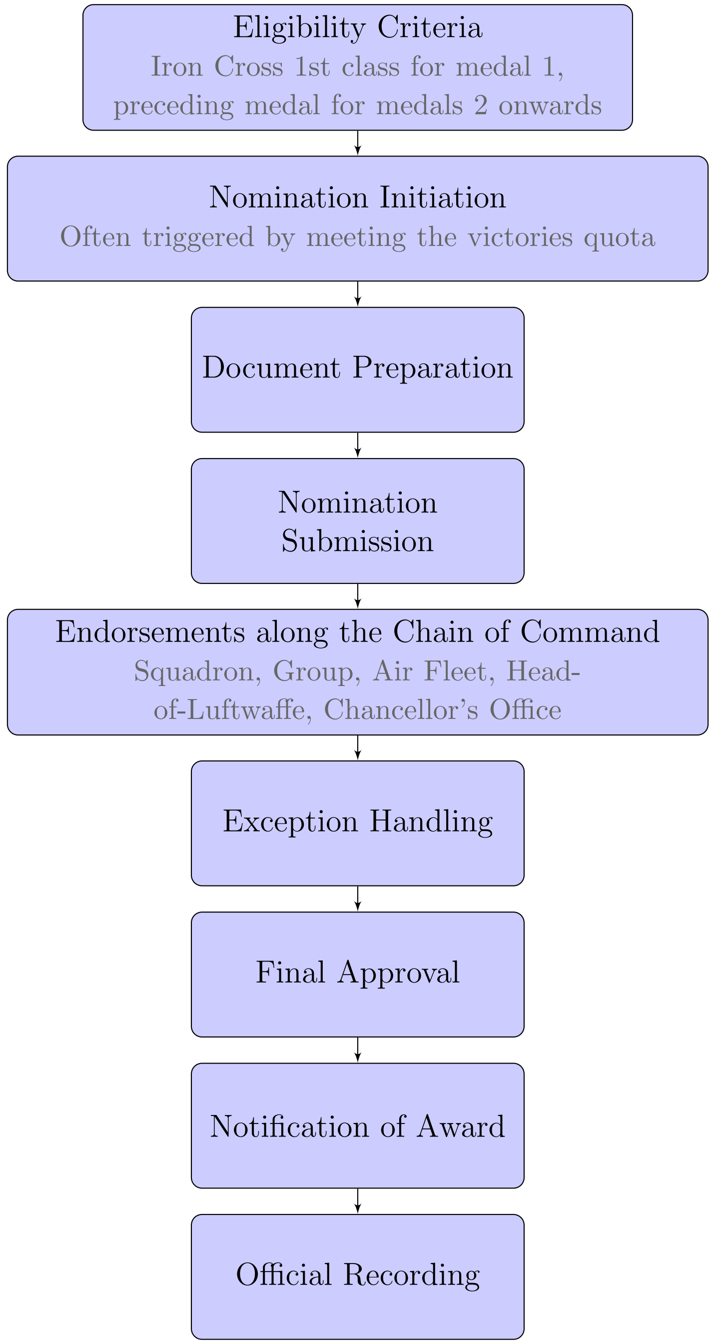


Notes: Figure shows the level of the victories quota over time and front for medals 1, 2 and 3. Dashed lines indicate the level on the Eastern front, solid lines on the Western front. Eastern front quotas are vertically offset to make them visible when they coincide with the Western front level.

Figure A2: The Knight's Cross medal

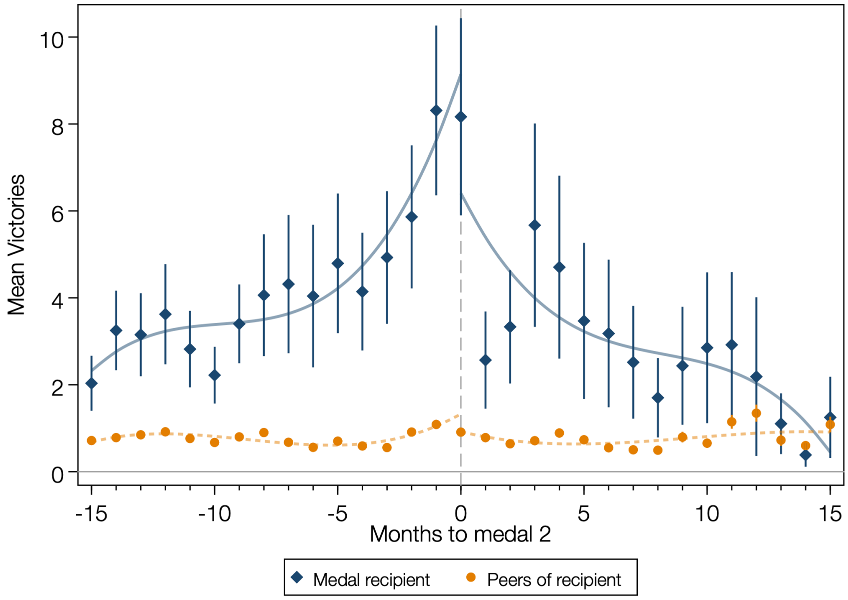
Notes: An original *Ritterkreuz des Eisernen Kreuzes* medal (medal 1). The most common variant of the medal was manufactured by Steinhauer & Lück and measured approximately 48 by 54 millimeters. We manually pixelate the swastika at the center of the medal, in line with German legal requirements.

Figure A3: Process of medal award

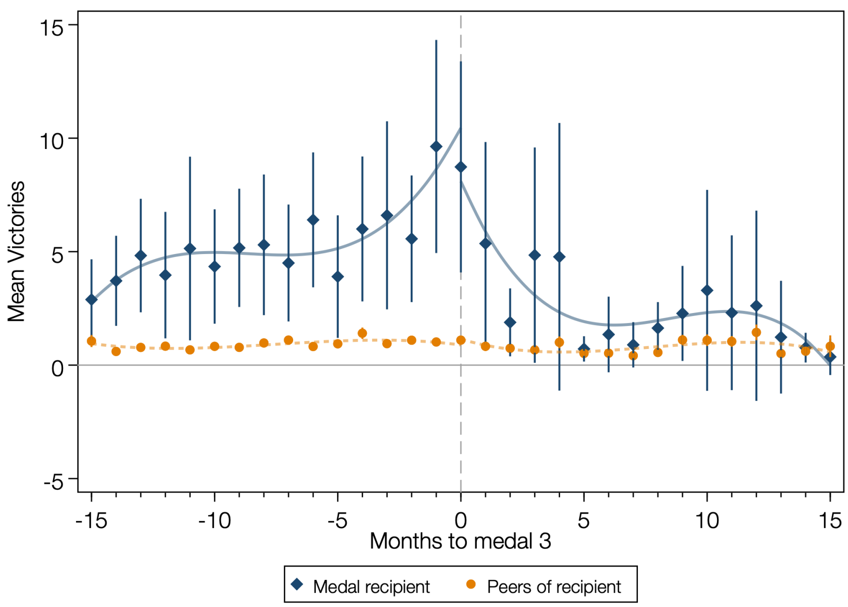


Notes: The process of qualification, nomination, and notification for the Knight's Cross medals, adapted from Williamson and Bujeiro (2004).

Figure A4: Raw victory rates over higher medal event time



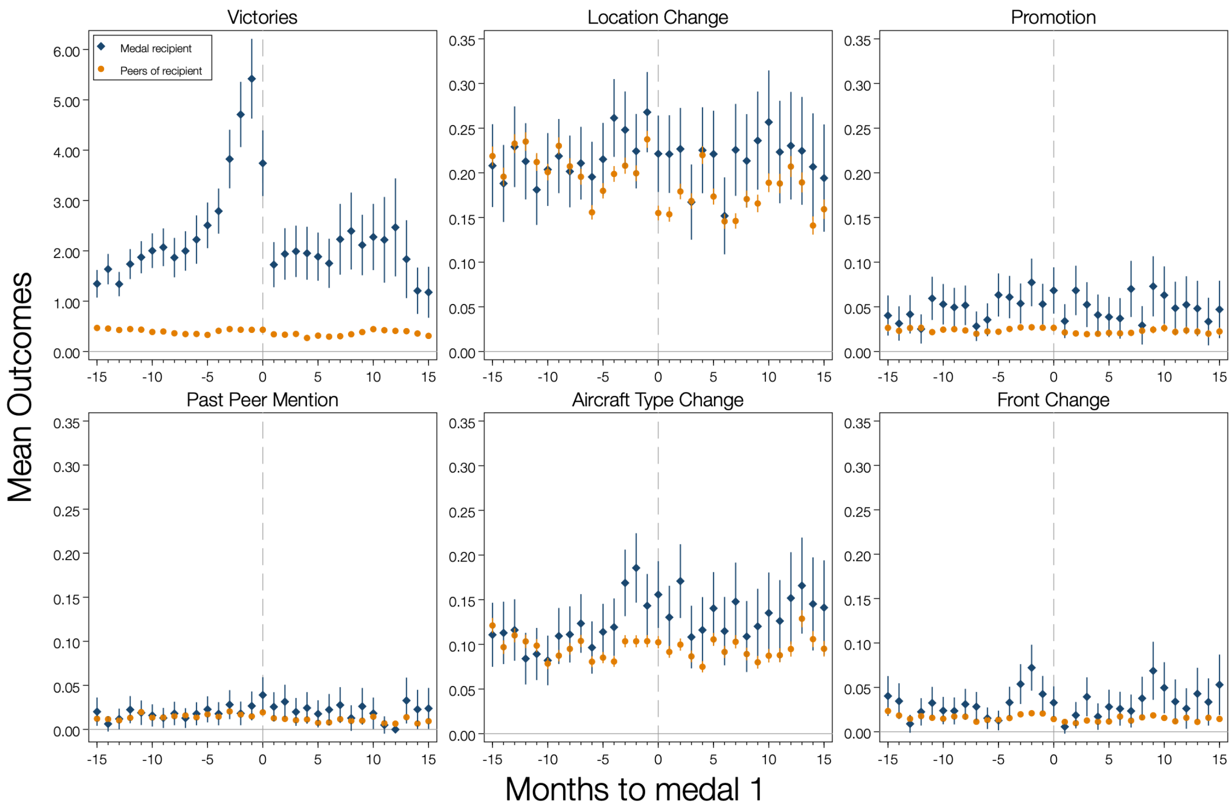
(a) Medal 2



(b) Medal 3

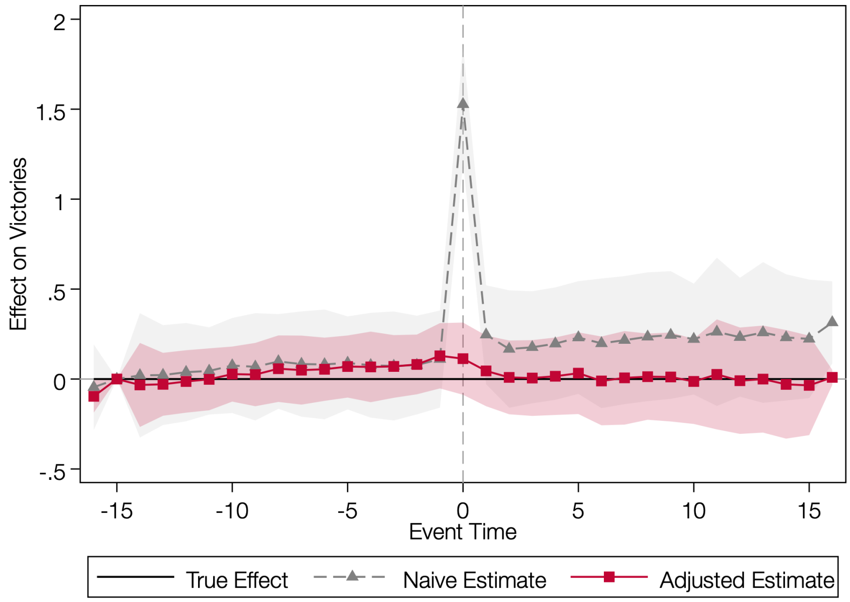
Notes: Figure shows mean monthly victories $\bar{Y}_j$ in blue over time relative to winning the medal 2 (upper panel) and medal 3 (lower panel). The orange line shows the mean monthly victories of peer pilots who fly at the same airbase and in the same month as the focal pilot. Bars give the 95% confidence intervals for the mean. Lines are splines fitted to cubic polynomials of the data.

Figure A5: Trends in other variables over medal event time

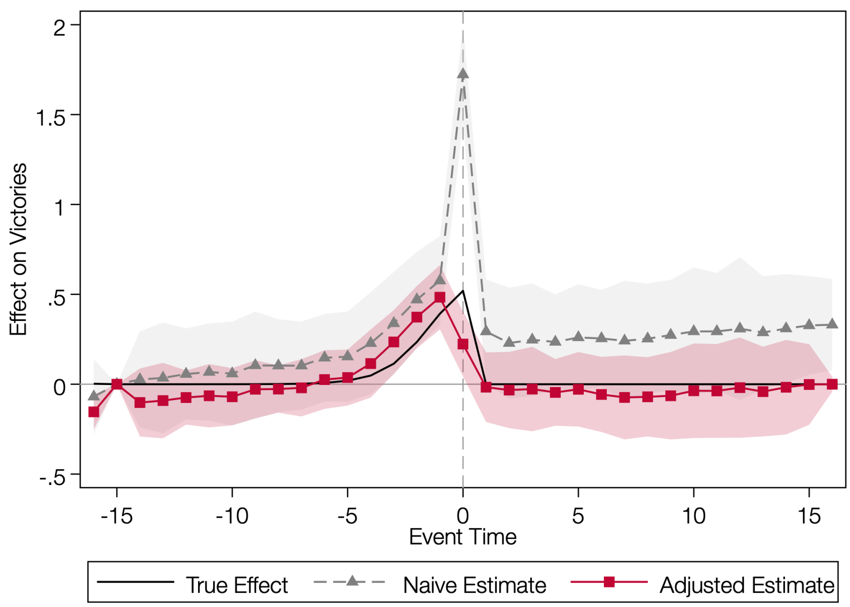


Notes: Each panel shows in blue the mean value of a variable over time relative to winning the medal 1. The orange line shows the mean value for peer pilots who fly at the same airbase and in the same month as the focal pilot. The variables (left-right, top-bottom) are the victory rate, and indicators for being assigned to a base more than 500km away, a promotion in rank, a former peer being mentioned in dispatches (the treatment in Ager et al. (2022)), a change of aircraft type, and a change in combat theater. Bars give the 95% confidence intervals for the mean.

Figure A6: Comparison of naive and adjusted estimators in synthetic data



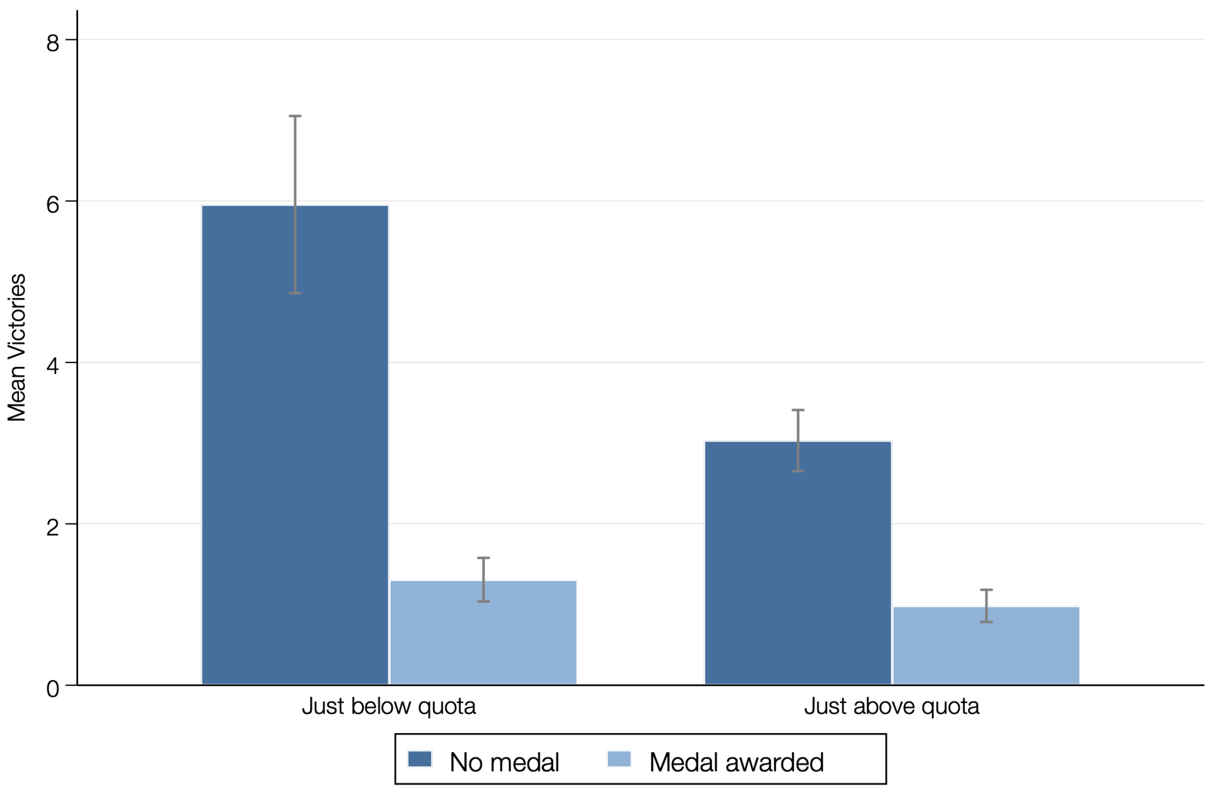
(a) Simulated effect under the null hypothesis ($\alpha = 0$)



(b) Simulated effect under the alternative hypothesis ($\alpha = 0.5$)

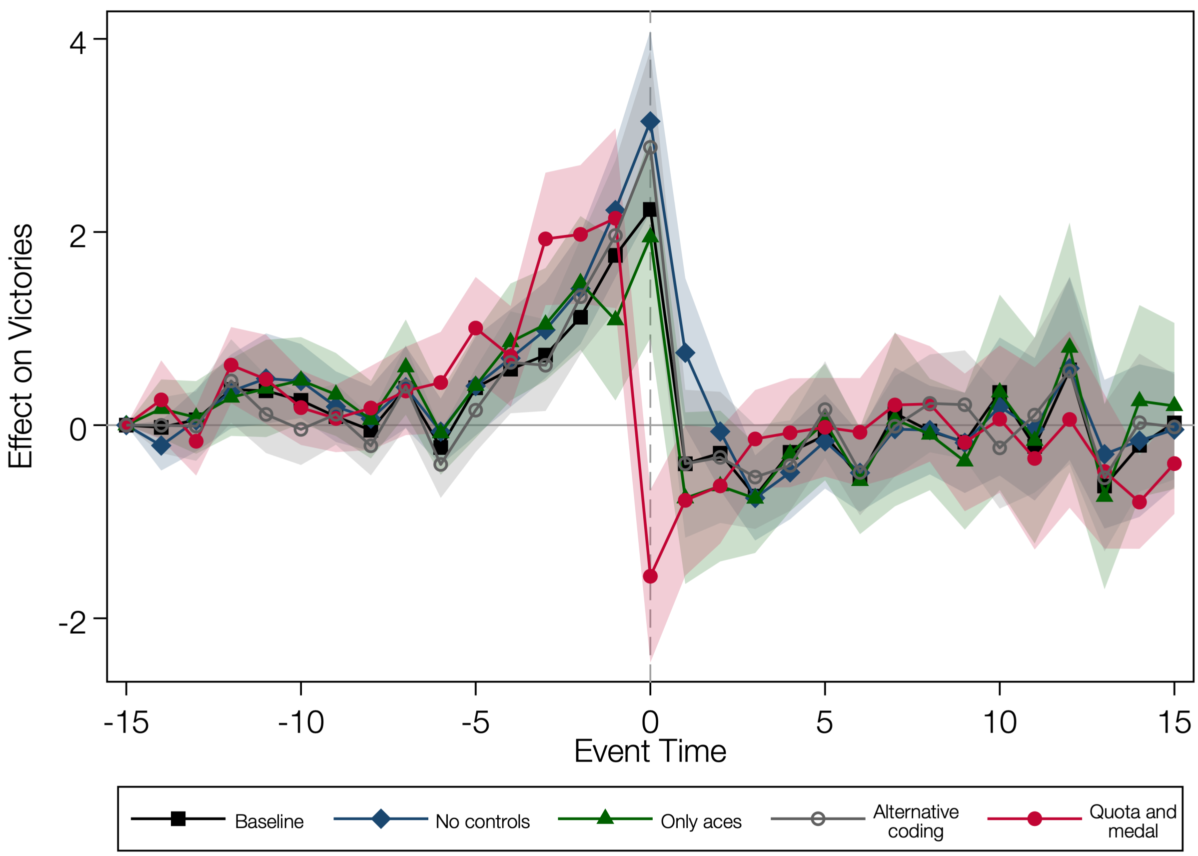
Notes: Figure shows the mean event study coefficients as estimated in 50 Monte Carlo simulations of the DGP described in Appendix C, where each simulation generates $p = 25$ placebo quotas per pilot for the mechanical-correction procedure. In the upper panel, the DGP sets the true medal anticipation effect to match the null hypothesis of no effect. In the lower panel, the true effect is set by a modified logistic function and shown in black. The dashed gray line shows the naive event estimation (2) without any adjustment. The solid red line gives the mechanical-correction procedure estimates. 95% percentile confidence intervals for Monte Carlo simulation coefficients are given in the shaded areas.

Figure A7: Victory rates around quota



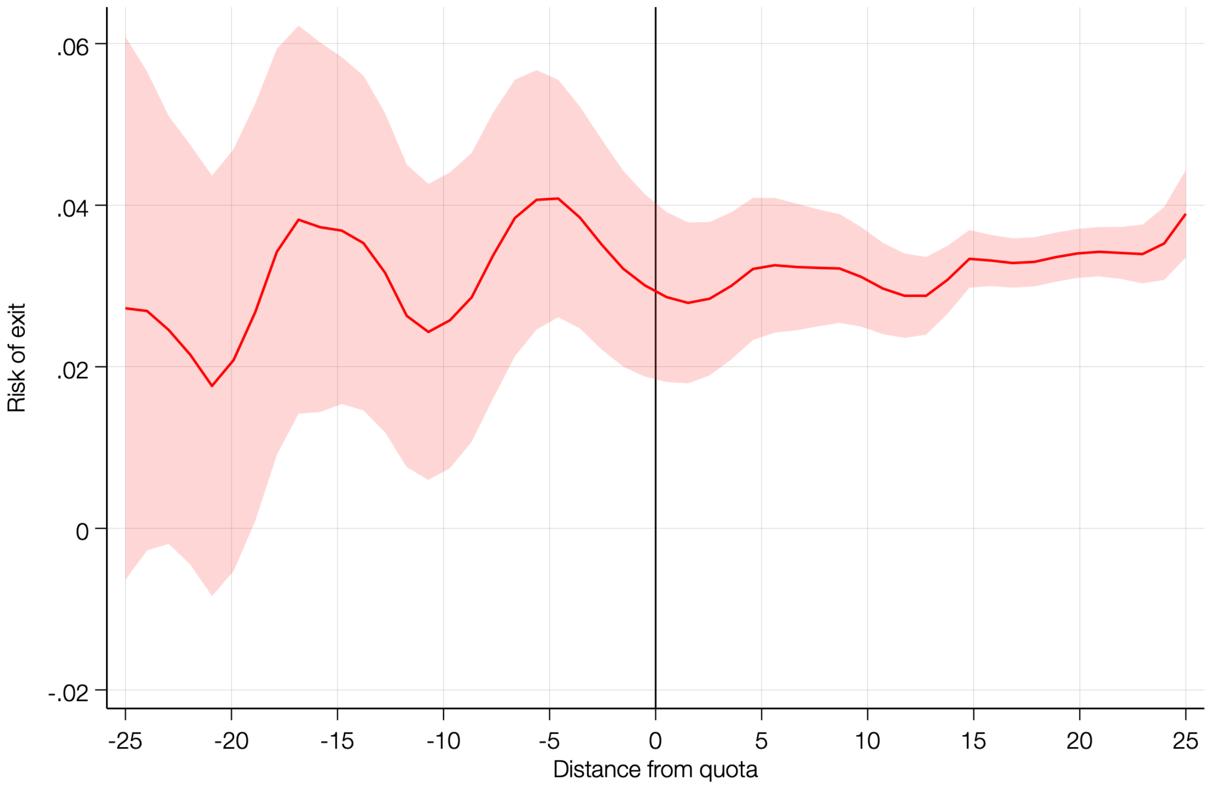
Notes: Each bar gives an average victory rate while within five victories of the medal quota. Light blue is for pilots who hold the medal in question, dark blue for those who do not yet have it. Both below and above the quota, award-holders exert less effort.

Figure A8: Pilot effort spikes around the medal quotas, robustness exercises



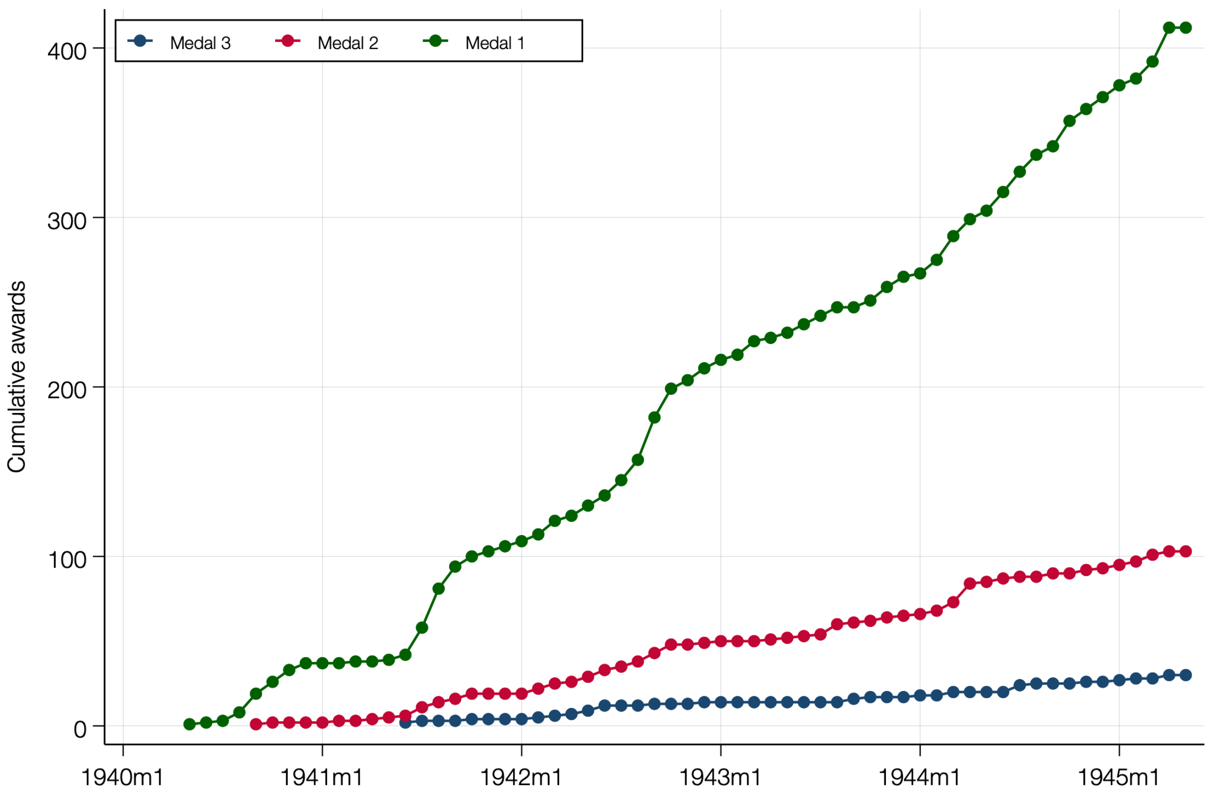
Notes: Estimated average treatment effects on victories per month as a function of their time relative to meeting the victories quota. The event study model (variants of (2)) includes 15 lags and leads of treatment; all further periods are absorbed into the 15th lag (lead), which is the reference period. The treatment occurs in period 0, so period 1 is the first period when pilots are fully treated. Specifications are explained in Section 2. Estimation via OLS with fixed effects for pilot and period. Standard errors clustered at the level of fighter squadrons. 95% confidence intervals for coefficient estimates are given in the shaded areas.

Figure A9: Pilot exit risk does not track quota distance



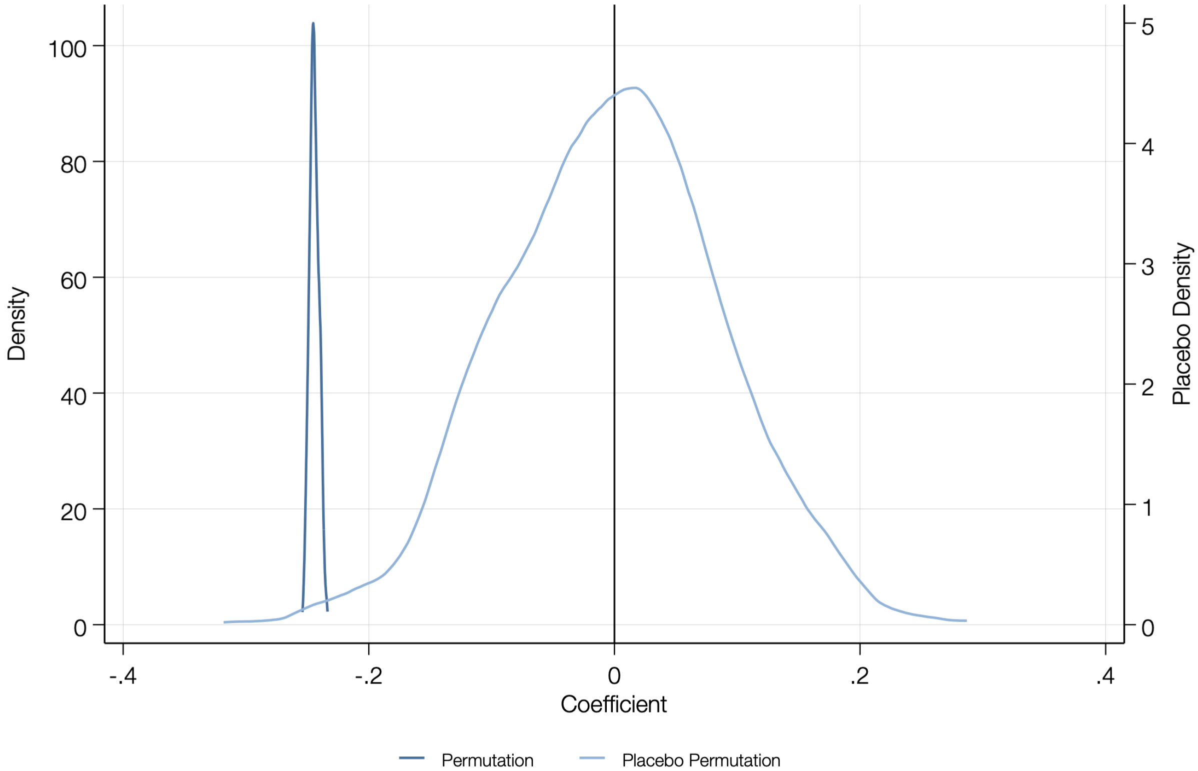
Notes: Figure shows the relationship between distance to the informal victories quota and the risk of exit. The horizontal axis shows the difference between a pilot's cumulative victories and the informal victories quota for the medal, which varies by period and front. The red line gives the local polynomial risk of exit conditional on distance, bounded by 95% confidence intervals in the shaded area.

Figure A10: The stock of medals awarded rises throughout the war



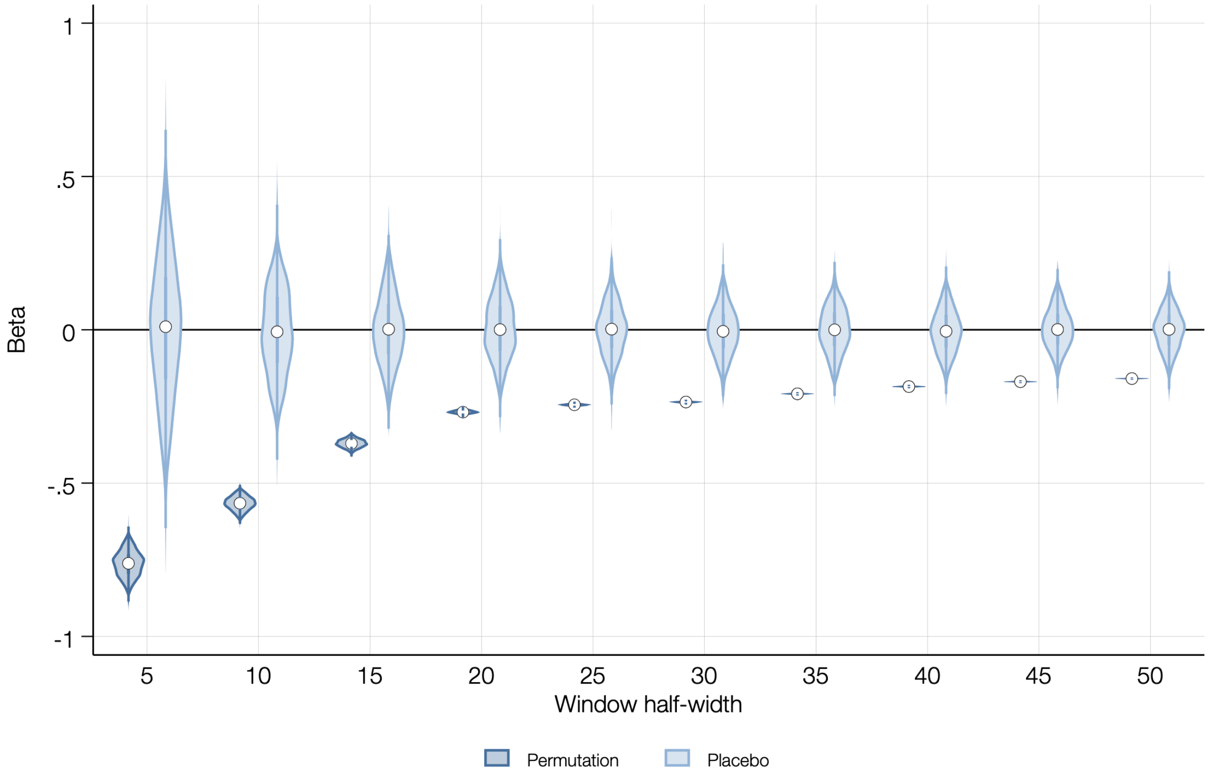
Notes: Figure shows the cumulative number of each type of medal awarded over the course of the war.

Figure A11: The estimated rate of cheapening is robust to permutations in medal order



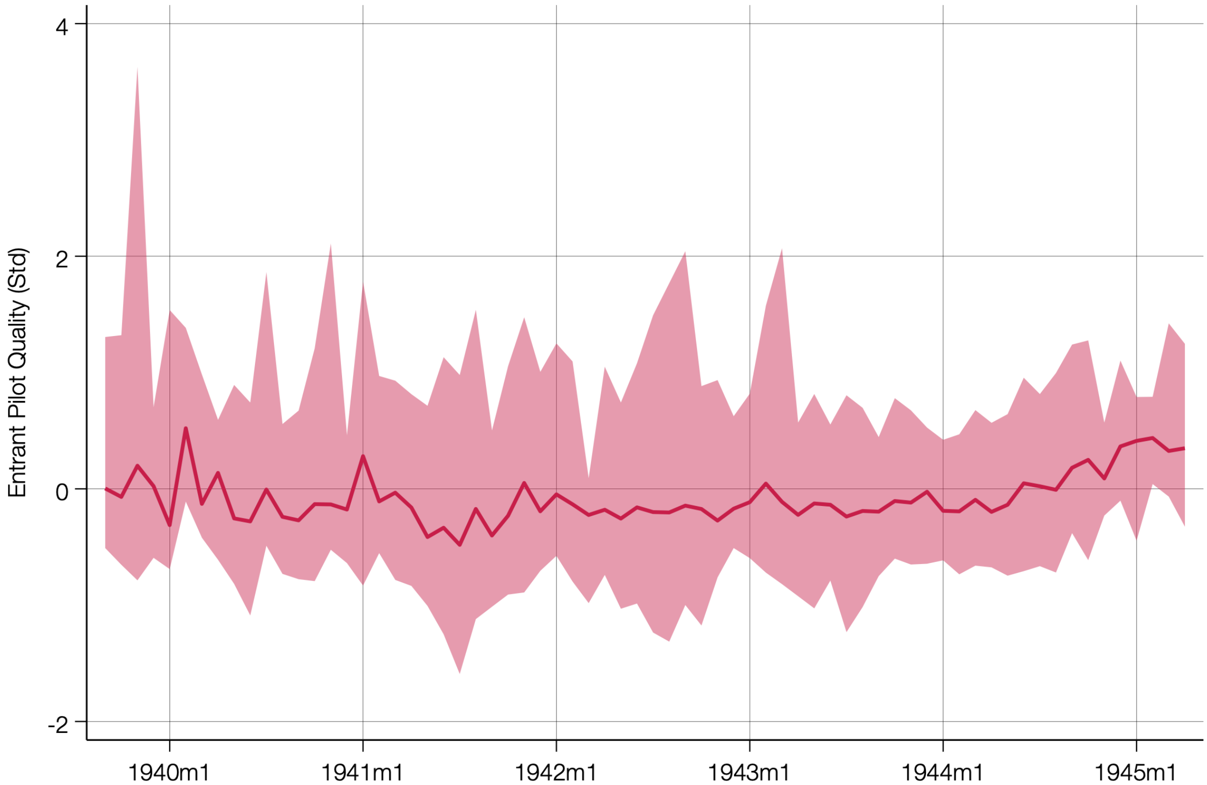
Notes: Figure shows the kernel density function of coefficient estimates of β in (3), permuted $1,000$ times to resolve the ambiguous order of same-day medalists. The darker line shows the distribution of the slopes in the permutation set, where the order is permuted only *within* shared award dates. The lighter line shows the distribution of the slopes from a placebo set, where the order is permuted fully at random *across* all dates.

Figure A12: The estimated rate of cheapening is robust to the width of the comparison



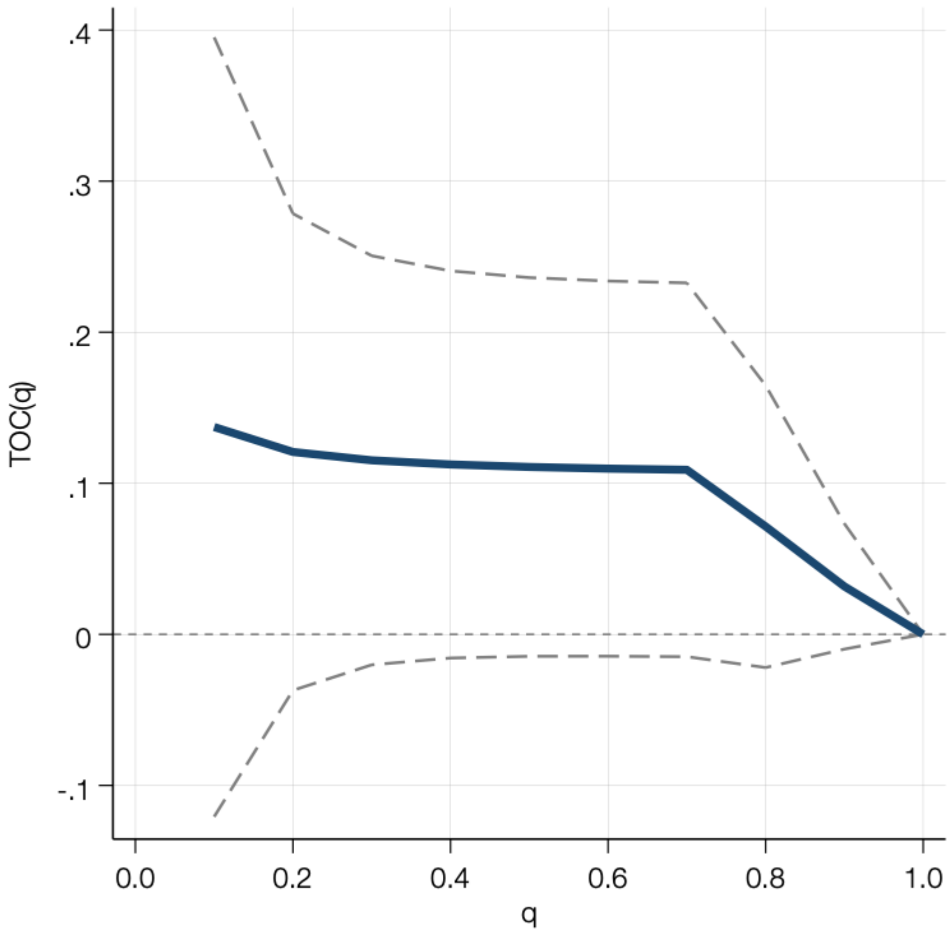
Notes: Figure shows the kernel density functions of coefficient estimates of β in (3) across a range of relative comparison window sizes w. Each distribution represents $1,000$ permutations. The darker distributions shows the slopes in the permutation sets, where the order is permuted only *within* shared award dates. The lighter distributions shows the slopes from placebo sets, where the order is permuted fully at random *across* all dates.

Figure A13: Ability of entrant pilots over time is constant



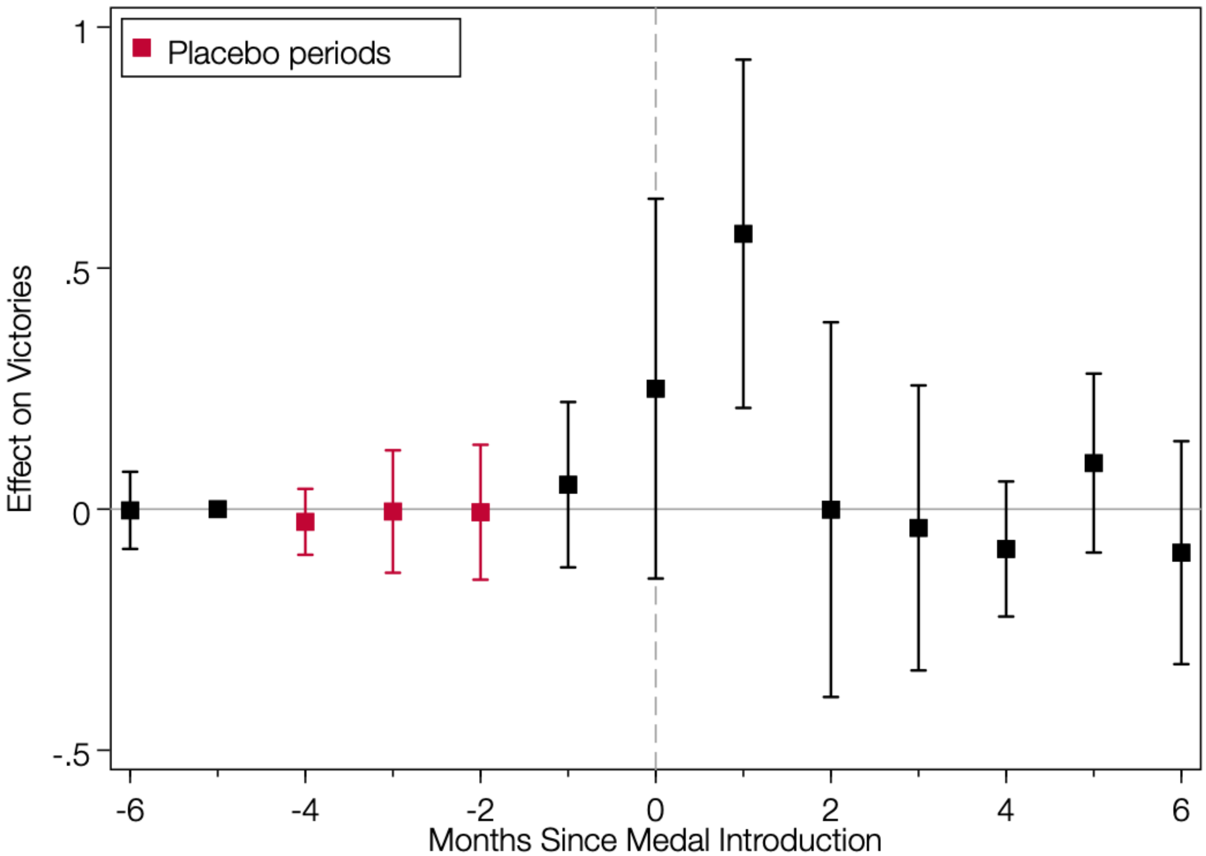
Notes: Figure shows the standardized ability (Mas and Moretti, 2009) of new pilots (in their first month as an active duty pilot) in each month of the war. The red line plots the median in each month. The shaded error spans the 10th to 90th percentiles of the data.

Figure A14: In zone treatment effects vary substantially across pilots



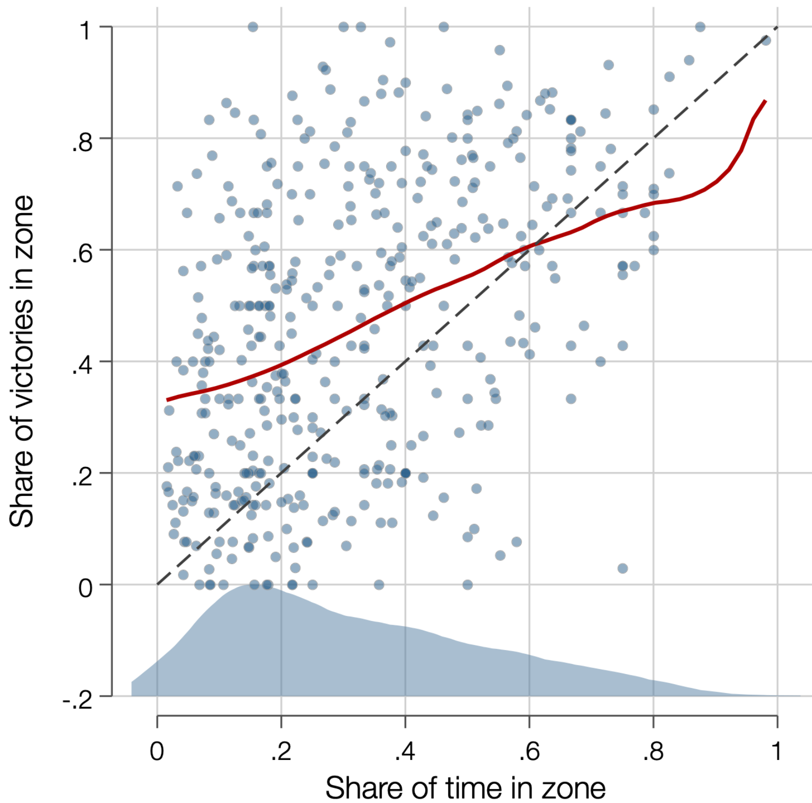
Notes: Figure shows the targeting operator characteristic (TOC) curve for the effect of being in zone of the medal quota. We estimate the effect of being in zone on the monthly victory rate, using noble birth and officer status as covariates X_i in the causal forest estimation framework of Yadlowsky et al. (2023). Treatment effect sizes are ordered from largest to smallest according to the prioritization score $S(X_i)$. We then compute $\text{TOC}(q) = E[Y_i(1) - Y_i(0) \mid S(X_i) \geq F^{-1}_{S(X_i)}(1 - q)] - E[Y_i(1) - Y_i(0)]$. The y-axis reports the increase in the ATE over unconditional ATE for the top q proportion of units. Under no treatment effect heterogeneity across X_i, the area under the curve (RATE) tends to zero.

Figure A15: Pilot effort increases when new medals are introduced: placebo



Notes: Figure shows the average treatment effect on the treated (ATT) of medal introductions on monthly victories for medal-sensitive pilots who are in the top quintile of conditional average treatment effects for the in zone treatment, using the remaining pilots as control units. Estimation via generalized synthetic control (Xu, 2017) with covariate matching on combat front and $r = 1$ latent factors. Periods indicated in red are excluded from the model fitting step as placebos. ATTs are re-centered relative to a reference period, $t = -5$. Bootstrap 95% percentile confidence intervals are indicated by the capped spikes.

Figure A16: Time 'in zone' contributes disproportionately to overall victory tally



Notes: Figure shows a scatter plot. Each point is a pilot who ever enters the zone of a medal (defined as the quota minus fifteen victories). The x-axis gives the share of *months* spent in zone for each pilot, and the y-axis gives the share of their total *victories* achieved while in zone. The red line gives the local polynomial fit of the data. The dashed line is the 45° line. The shaded area represents the histogram of the x-axis variable.

B Data Appendix

Ability estimation

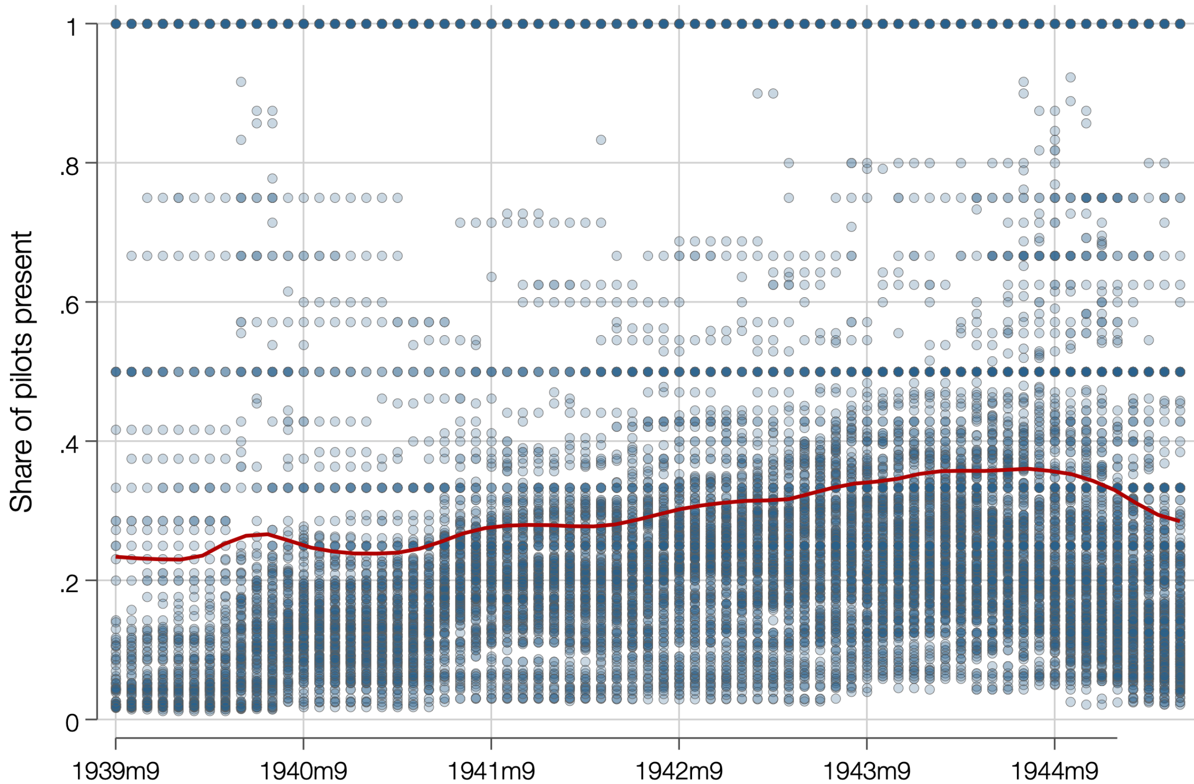
Our procedure for estimating the ability of pilots, given the presence of peer effects, is the empirical model of Mas and Moretti (2009):

$$Y_{i,t,c,s} = \theta_i + \mathbf{M}'\vartheta_{Cit} + \pi N_b + \gamma R_{t,f} + \varepsilon_{i,t,c,s} \tag{5}$$

The monthly victory tally of pilot i at time t in squadron c with peer set s is $Y_{i,t,c,s}$. We interpret θ_i—the individual permanent productivity—as the inherent ability of pilots. Our goal is to recover this coefficient vector. ϑ_{Ci} is a vector denoting each specific combination of peers alongside the focal worker in each period. N_b and $R_{t,f}$ capture airbase and front-by-time specific shocks, respectively. The idiosyncratic error term is $\varepsilon_{i,t,c,s}$.

Identification comes from observing different peer compositions over time. We show that, in any given month, squadrons typically only contain a fraction of the total pilots who ever serve in the squadron in Figure A17. Squadron entry was as-good-as-randomly assigned, driven by operational needs (Caldwell, 2012). Most squadron exits are death or permanent incapacitation, there is a large stochastic component to the composition changes. We define peers as those fellow pilots flying in the same squadron and month.

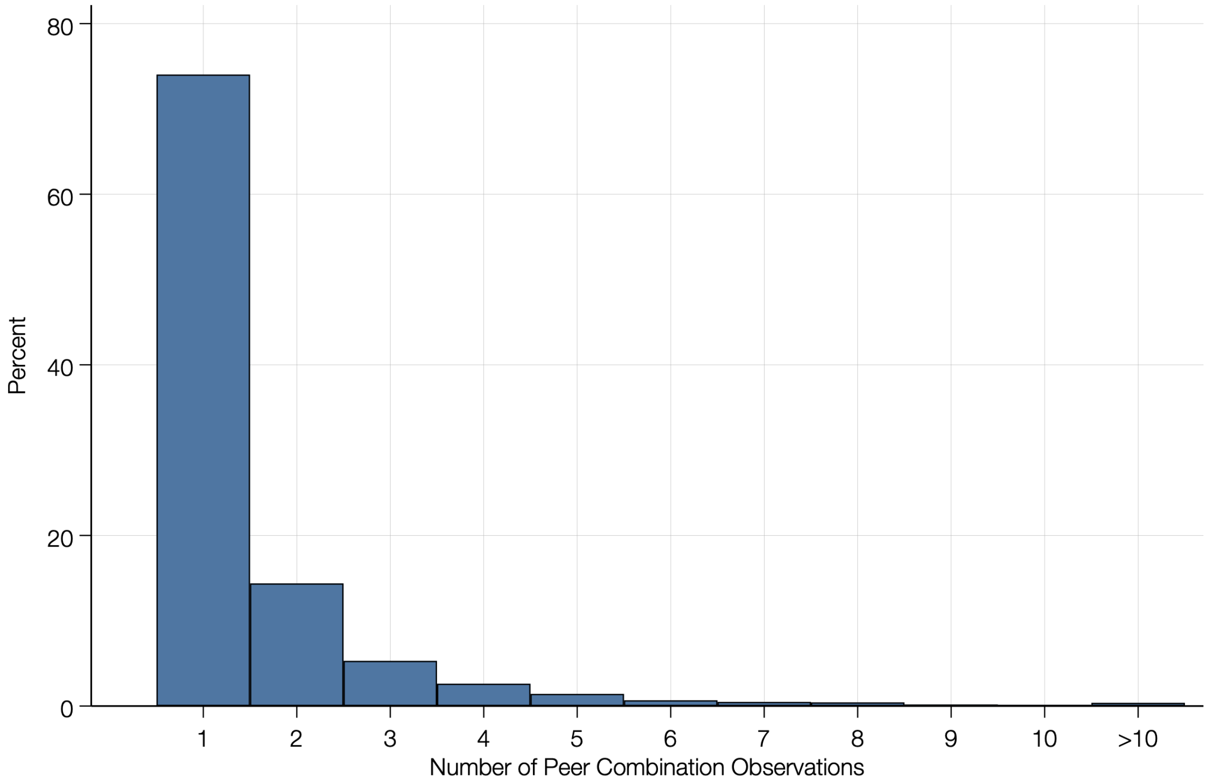
Figure A17: Pilot ability estimation: squadron composition variability



Notes: Each point is a squadron-month. The y-axis position is the number of active pilots in the squadron divided by the total number of pilots ever observed serving in this squadron. The x-axis is the monthly time. The red line is a local polynomial fit of the data.

Among 82,348 monthly pilot observations, we observe 50,075 different peer combinations. We show the distribution of the number of appearances of each peer combination in Figure A18. Since the ratio of peer combinations to total observations is high, we cannot estimate the full set of θ_i—individual fixed effects capturing permanent productivity. We adopt the procedure from Hamilton et al. (2022): study the first-stage model fit when we exclude peer combination fixed effects that appear fewer than δ times over a range of δ. As δ increases, we can estimate θ_i for more pilots, but we will sacrifice model fit if specific peer-combination effects are essential to the nature of the work.

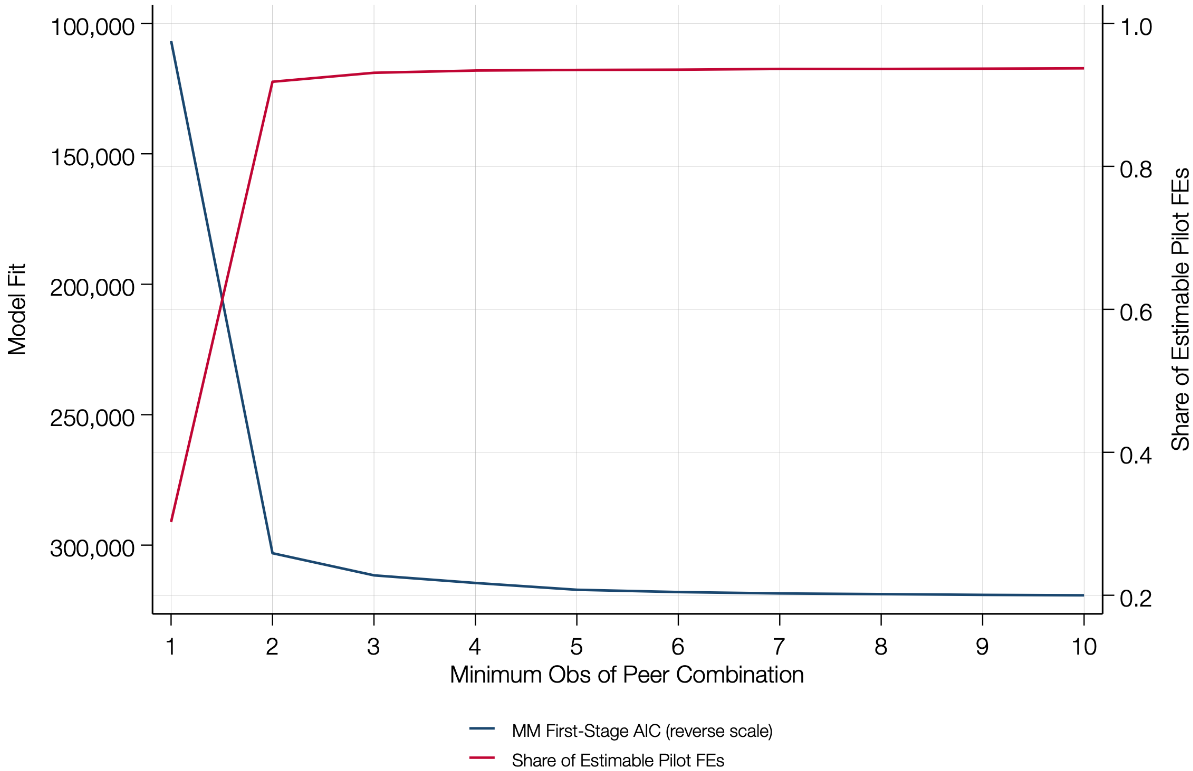
Figure A18: Pilot ability estimation: peer combination frequency



Notes: Each observation is a specific combination of peers. We plot the histogram showing the number of times we observe this combination in the monthly panel. The median peer combination appears only once.

We choose the δ representing the best trade-off between fit and coverage. This trade-off is plotted in Figure A19—we choose $\delta = 2$ at the "elbow". This allows us to estimate the ability of most pilots while maintaining a strong model fit.

Figure A19: Pilot ability estimation: peer fixed effects and model fit



Notes: We plot the trade-off between model fit and estimable permanent productivity under the procedure from Hamilton et al. (2022). The x-axis is δ, the number of observations for a specific combination of peers. We successively examine fixed effects for peer combinations we observe less than δ times and refit the model. We report model fit using the Akaike Information Criterion (AIC), where a lower value represents a better model fit. With a stronger δ restriction, we can estimate a successively greater number of individual pilot permanent productivities.

In Table A6, we compare the attributes of the pilots whose θ_i we can and cannot estimate under this choice of δ. As expected, estimable pilots serve for more extended periods and change squadrons more frequently, but are also higher performers.

Table A6: Balance of pilot attributes based on inclusion in ability estimation

	Estimable	Non-Estimable	T-test Difference
Mean Monthly Victories	0.863	0.788	0.075
	(1.040)	(0.745)	
Mean Front	0.332	0.344	-0.012
	(0.438)	(0.471)	
KCR	0.088	0.014	0.074***
	(0.284)	(0.119)	
Total Victories	11.483	3.065	8.418***
	(23.842)	(5.184)	
Tenure	16.926	8.165	8.761***
	(17.765)	(11.843)	
# Squad Changes	0.671	0.127	0.544***
	(1.248)	(0.564)	
N	4,664	417	5,081

Notes: Characteristics of pilots who we can and cannot estimate individual permanent productivity for using the modified Mas and Moretti (2009) procedure, dropping peer-combination-specific fixed effects that appear only once. T-test for the difference in means between groups.

C Monte-Carlo Validation of the Medal Anticipation Effects Estimator

Here we describe the details of the Monte-Carlo test we implement to demonstrate the validity of our estimation procedure for recovering medal anticipation effects, as introduced in Section 2.

1) **Data-generating process.** We simulate $N = 400$ pilots and $T = 70$ monthly periods. Each pilot draws an underlying ability $a_i \sim \log \mathcal{N}(0, 0.4^2)$ and a common quota $q_i = 20$. Baseline victories follow a Poisson process with a mean $\mathbb{E}[Y_{it}^{\text{base}}] = \exp(a_i) + \varepsilon_{it}$.

2) **Medal anticipation effects.** Extra effort in anticipation of the medal θ_j is added via a modified logistic function over distance to the medal quota d_{it}:

$$Y_{it}^{\text{extra}} = \begin{cases} e(d_{it}) & d_{it} > 0 \\ 0 & d_{it} \leq 0 \end{cases} \tag{6}$$

$$\text{where } e(d_{it}) = \frac{\alpha}{[1 + \exp(\beta(d_{it} - \delta))]\,(1 + \Gamma \cdot a_i)} \tag{7}$$

   We choose $\{\alpha, \beta, \delta\}$ to roughly match the shape in the raw data in Figure 2. The parameter Γ introduces a correlation between ability and the size of the anticipation effect, which we set to be weakly positive. Under the null hypothesis of no medal anticipation effect, we set $\alpha = 0$. Total victories are simply $Y_{it} = Y_{it}^{\text{base}} + Y_{it}^{\text{extra}}$.

3) **Exits.** Pilots are killed in a given period if they lie above the $(1 - \kappa)$ percentile of the risk distribution. Risk is given by:

$$r_{it} = y_{it} \cdot \Sigma \cdot (1 + \Phi(1 - a_i)) - \Omega \cdot a_i + \varphi_{it} \tag{8}$$

   Where y_{it} is Y_{it} standardized to have mean zero and sd one in each period, $\Sigma \geq 0$ is the elasticity of risk with respect to victories, $\Phi \geq 0$ weakens this elasticity for high-ability pilots, and $\Omega \geq$ lowers the baseline risk for high-ability pilots. $\varphi_{it} \sim \mathcal{N}(0, 1)$ is idiosyncratic risk. We roughly match our data by setting $\kappa = 0.05$, so 5% of pilots die each period.

4) **Placebo quotas.** Each pilot draws a placebo quota from a uniform integer distribution on $[\underline{\mu} \cdot q_{it}, \overline{\mu} \cdot q_{it}]$. We set $\underline{\mu} = 0.1$ and $\overline{\mu} = 2.0$. The placebo quota is drawn $p = 25$ times.

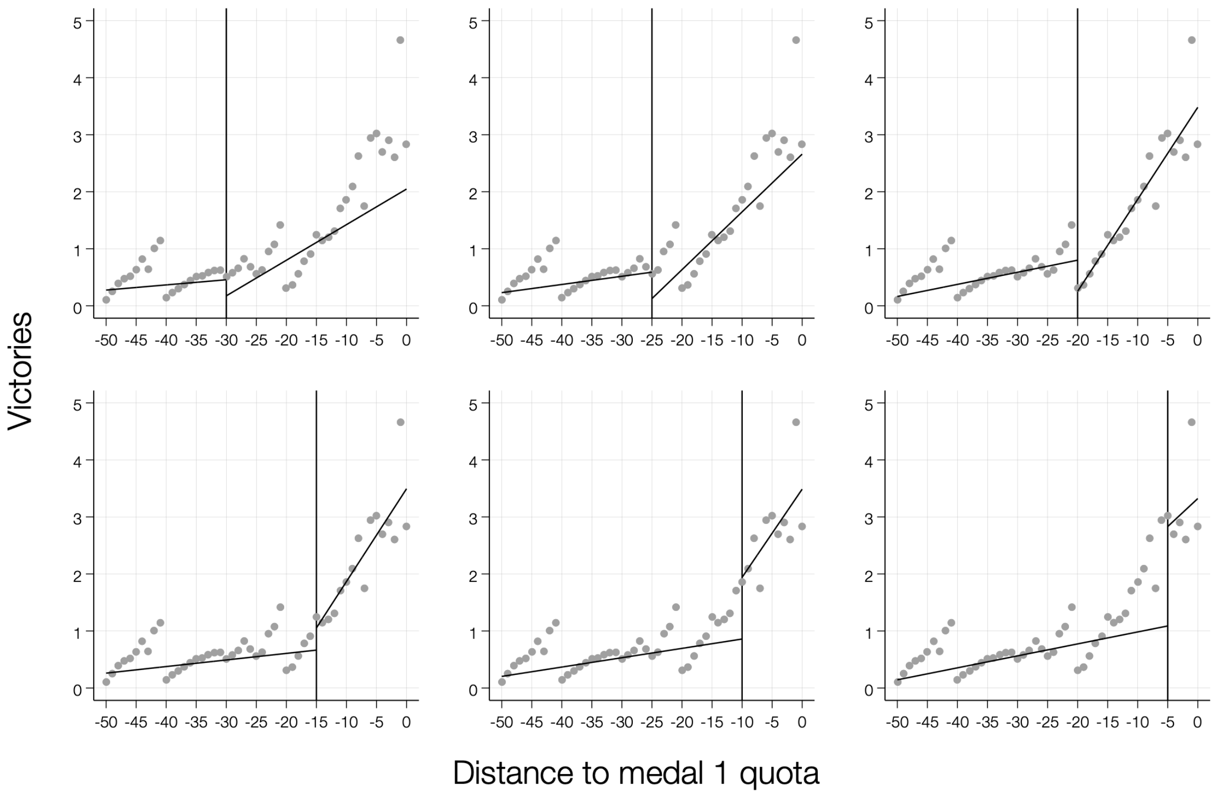
Figure A6 visualizes the results of 50 Monte Carlo repetitions, where each repetition generates $p = 25$ placebo quotas per pilot. The dashed gray line shows the biased naive estimator corresponding to $\{\hat{\theta}_j^{\text{naive}}\}$: it overshoots the true pre-quota effort by roughly 40 percent and exhibits a spurious positive effect *after* the quota has been reached. By contrast, the solid red line (placebo-adjusted $\{\hat{\theta}_k^{\text{adj}}\}$) tracks the black truth almost perfectly, and its confidence band covers θ_j at every horizon, under both the null and alternative hypotheses. The simulation, therefore, confirms that the adjusted estimator removes the mechanical bias ξ_j.

D Instrumental Variables Approach to Estimation

This section describes our approach to estimating the in zone effect that relies on a conventional instrumental variables estimator.

To capture the effects of quota proximity in a static regression, we define a dummy for being in the zone of the medal quota as being up to D victories below the quota Q and not holding the medal. We calibrate D by estimating regression discontinuity analysis around a series of candidate entry points into the zone, $D \in \{-30, -25, ..., -5\}$. We show in Figure A20 that the slope discontinuity in the victory rate, estimated using (Calonico et al., 2015), is most prominent at $D = -15$.

Figure A20: Local polynomial estimates of zone entry

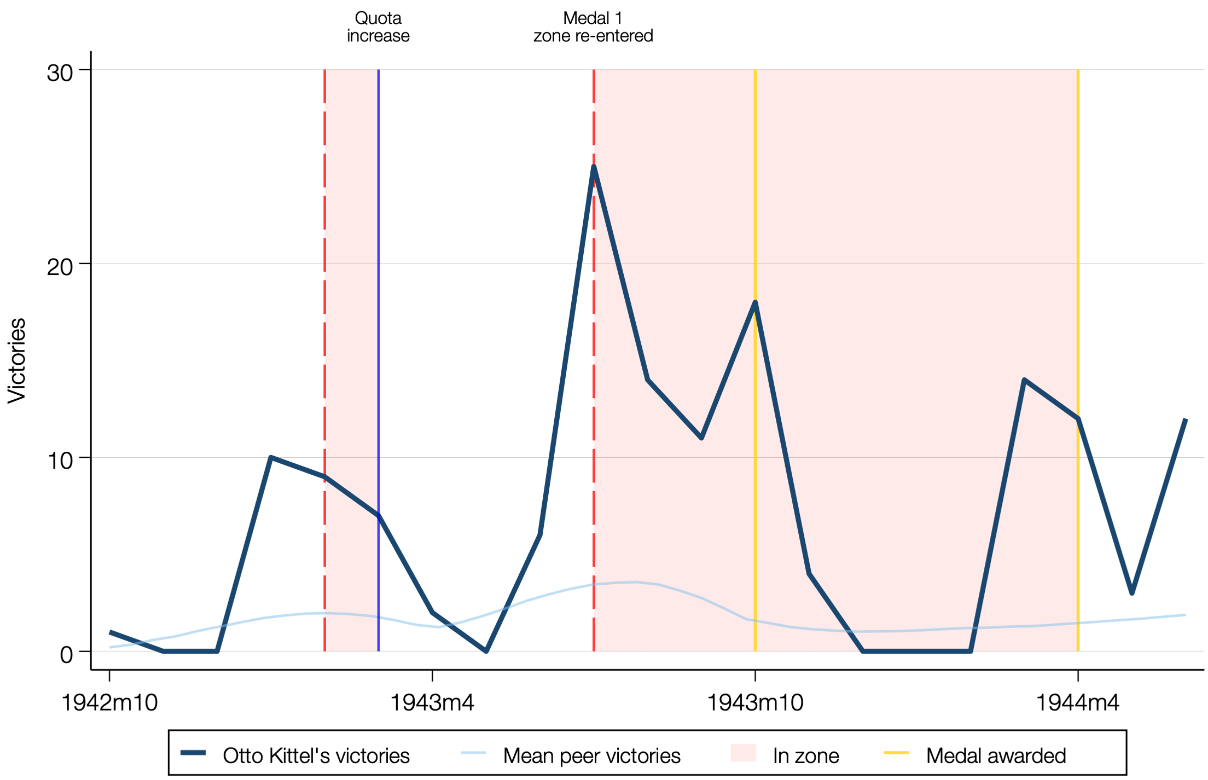


Notes: Figure shows a series of regression discontinuity plots estimated around candidate "in zone" entry points, the stock of pilot victories is distance $D \in [-30, -25, ..., -5]$ below the informal victories quota for medal 1. Using the method of Calonico et al. (2015), the data is partitioned into bins, where the number of bins is selected using polynomial estimators to mimic the variance evenly-spaced method. The local linear regression has a uniform kernel. The dependent variable is the number of victories for a given pilot in a given month.

Our instrument is a dummy indicating a quota increase relative to last month for the medal. When the quota is increased (decreased), pilots are moved out of (into) the zone even as their victory tally is constant. Given the relative abundance of quota increases in our data, increases are the most likely of all the changes to push the marginal pilot *away* from the zone of the medal. Hence, we expect our first-stage and reduced-form regressions to give a negative coefficient on the instrument.

We illustrate the causal mechanism described above with the case of Otto Kittel (Figure A21). Kittel is among the highest-scoring aces in our dataset, with over 250 victories at the time of his death in combat in 1945. In early 1943, Kittel reached 40 victories: enough to be within just 10 victories of the medal 1 quota on the Eastern Front. We can see his performance jumping up the month before already, when he is at a distance of 20 victories to the quota, and scores 10 in a single month. However, just as he closed in on the quota in March 1943, the required number of victories was increased from 50 to 75, putting medal 1 far beyond Kittel's immediate prospects. Having been pushed out of the zone by Luftwaffe policy changes, his performance slumped for several months (despite his squadron having some success). Eventually, he closes in on the new quota. He enters the zone in July and continues to perform strongly in August and September. Finally, he receives medal 1 in October, by which time he is already in the zone for the next medal. We see that in the period after the quota increase, performance remains low, since the medal quota is distant. As the Kittel becomes in zone again, performance increases dramatically. Our empirical results will use the quota variation to show that these examples are indeed indicative of the general pattern.

Figure A21: Quota changes and performance: Case study evidence



Notes: Figure shows the performance of fighter ace Otto Kittel (b. 1917) over time. The navy line gives the monthly victory score. The light blue line gives the mean score of squadron peers. The red vertical lines indicate when the pilot is in zone (within fifteen victories) of a medal quota. The gold lines indicate when a medal was awarded. The bold blue line in the lower panel indicates when the quota for medal 1 increased, pulling Kittel out of the zone.

The instrument's relevance is guided by the mechanical connection between the quota level Q and the treatment $\mathbb{1}(V - Q < D)$, where our cutoff $D = -15$. The exclusion restriction requires that the quota changes only affect pilot performance by changing proximity to the medal quota. One might be concerned that the same quota changes affect multiple pilots who compete with their peers to achieve the medal first (Ager et al., 2022). First, the variation in Q will only affect pilots with many victories V close to Q. Figure 3 shows that few pilots are in this range of victories. Second, we can explicitly exclude squadrons where this is a concern. In our preferred specification, we exclude squadrons with multiple treated pilots to preclude our estimation from capturing competitive rather than motivational responses to the quota change. Moreover, the remainder of the analysis in this section uses a combined in zone treatment; a pilot is in zone if they are in zone of any of the first three medals. The relevant quota change is the one for the medal they are currently pursuing. This allows us to use more of the variation in quota levels and reduces the likelihood that a quota change spills over across pilots (e.g., on the same airbase).

In Table A5, we first estimate the static equivalent of the analysis in Figure 5. While the point estimates in columns 1–2 are consistent with a positive in zone effect, the magnitudes are significantly smaller than suggested by the event study. In column 3, we show the reduced form effect of the quota increase on the victory rate. The effect is significant and negative—pilots forced away from the medal by exogenous changes in quota are demotivated by the medal's increased distance from their current score. We then present 2SLS estimates in columns 4–6. In column 4, we show a baseline specification, absent controls or the sample restriction for multiple in zone pilots. In columns 5 and 6, we restrict the sample to squadrons with at most one in zone pilot, to remove the potential for competitive responses. The final column shows the preferred specification, adding covariates corresponding to the event study specification in (2). The Kleibergen-Paap rk Wald F statistic and Anderson-Rubin p-values indicate our results are suitably powered and statistically significant in all three 2SLS specifications. Overall, we find effects consistent with, though larger than, our preferred event study approach.