

Never Enough: Dynamic Status Incentives and the Rat Race*

Leonardo Bursztyn
Ewan Rawcliffe
Hans-Joachim Voth

October 2026

Abstract

How can people be motivated to provide sustained or repeated effort? We study hierarchical status incentives as a way to create a “rat race”. Using data on more than 5,000 German air force pilots in World War II, we show that performance rises sharply as pilots approach medal eligibility, then falls after receiving the medal. The same pattern recurs as pilots near higher-prestige awards, and status-conscious pilots respond strongly to newly introduced medals. Medals also lose cachet as lower-ability pilots receive them, making new awards more valuable. A tiered, expanding status system can therefore repeatedly harness status concerns to extract effort.

JEL: D22, D91, M55, N44

*Bursztyn: University of Chicago and NBER, bursztyn@uchicago.edu; Rawcliffe: Harvard University, ewanrawcliffe@g.harvard.edu; Voth: University of Zurich and CEPR, voth@econ.uzh.ch. We thank Dmitry Arkhangelsky, Nava Ashraf, Davide Cantoni, Zoë Cullen, Georgy Egorov, Ruben Enikolopov, Emir Kamenica, Nancy Qian, Muriel Niederle, Ricardo Perez-Truglia, Dominik Rohner, Noam Yuchtman and Katia Zhuravskaya for helpful comments and conference participants at the University of Chicago and the Geoeconomics Conference in Villars for thoughtful suggestions. Utkarsh Dandanayak and Elias Hadj Ammar provided outstanding research assistance.

*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

Non-financial rewards—medals, awards, and prizes—exploit people’s desire for status and their tendency to compare themselves to others (Frey and Gallus, 2017; Perez-Truglia, 2020; Cullen and Perez-Truglia, 2022). In static settings, there is both empirical evidence and good theoretical reason why such rewards should elicit effort (Besley and Ghatak, 2008; Moldovanu et al., 2007; Kosfeld and Neckermann, 2011; Ashraf et al., 2014; Gallus, 2017). However, once a symbolic reward is received, motivation can fade. Unlike other incentives, status is inherently *relative*. A dynamic system of incentives, aimed to create a never-ending cycle of goal chasing, can potentially keep incentives sharp at low cost.

Many firms exploit status concerns through tiered product hierarchies when it comes to consumption: airlines, car rental companies, casinos, and credit card providers all offer silver, gold, and platinum tiers (Bursztyn et al., 2018). Pesendorfer (1995) formalizes the logic in a model of fashion cycles: a monopolist repeatedly introduces new product vintages; high-status consumers distinguish themselves by adopting the latest, while older vintages lose their signaling value as they diffuse to lower types. At the same time, status-based incentives may prove effective only over short horizons, or only when alternative status sources are limited.¹ Whether rat race strategies can create *sustained* worker effort in practice is unclear.

In this paper, we exploit highly granular data on how individuals respond to such an incentive structure. We evaluate the effect of creating a rat race on effort in the context of aerial combat during WWII. The setting has several advantages: stakes are life-or-death, output is well-measured, motivation is key, and pilots have high autonomy in choosing their level of effort (Ager et al., 2022). As is common in military and other bureaucratic settings, there is limited

¹The 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”.

leeway to offer instrumental benefits, so the armed forces 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*) are a particularly useful setting for our purposes. Outnumbered and outgunned for most of the war, they emphasized individual motivation (van Creveld, 1982). They engineered a sophisticated system of such incentives: the Knight’s Cross medals, “the most sought-after military decoration” (Obermaier, 1966, p. 21).² The first Knight’s Cross was introduced at the start of the war as the top award. As more and more soldiers reached this tier, ever more exclusive variants were added—a “ladder” of up to five medals. Such a system has the potential to motivate men who otherwise had few opportunities to distinguish themselves further. It could also backfire, creating jadedness about the repeated devaluation of (formerly) top awards. How these opposing psychological forces play out is an empirical question.

Section 1 presents our dataset, which extends Ager et al. (2022). That paper analyzes personal rivalries between pilots using mentions in dispatches—a highly prized but fleeting and unpredictable form of recognition, well-suited to studying spillovers from rivalry but a poor source of direct motivation.³ In contrast, we demonstrate the *direct* effect of a policy engineered to elicit effort through a highly prestigious award, distributed via a mix of objective criteria and discretion.

Section 2 studies the effects of the first medal—the lowest rung of the ladder. Eligibility was governed by informal, loosely enforced victory *quotas*, periodically changed throughout the war. Because a spike in performance mechanically precedes any performance-based award (Frey and Gallus, 2017), we employ an event study over time relative to meeting the medal quota.

Our key estimand is the causal effect of the incentive system on pilot effort. Effort is not directly observable; we measure output as aerial victories, focus on within-pilot variation, and residualize output against rich fixed effects for squadron, combat front, and calendar month—absorbing pilot and squadron quality, the intensity of combat, and the overall state of the war. Our estimand is the quota-timed component of this residual. Its interpretation

²The German armed forces, including the Luftwaffe, were involved in German war crimes during WWII. None of our results should be seen as endorsing a genocidal regime.

³They examine the victory and exit rates of former squadron peers when a pilot receives a mention in dispatches.

as effort requires only that luck, omitted variables, and measurement error are not systematically correlated with the timing of the quota event. We address the main challenge, joint production with peer pilots, with direct descriptive evidence and a Shapley decomposition exercise. For a subsample of pilots with surviving flight logs, we also observe behavior itself: pilots fly more combat sorties, and score more victories per sortie, as the quota approaches.

Why should effort rise near the quota? Scoring an aerial victory always required risk-taking. But pilots knew how close they were, and proximity promised a discrete jump in fame and recognition on top of the expected benefits of each marginal victory. In a standard elimination tournament (Rosen, 1986), the option value of prizes further up the ladder sustains incentives, so effort should not fall once a rung is reached. The goal-gradient hypothesis (Hull, 1934; Heath et al., 1999; Kivetz et al., 2006), quasi-hyperbolic discounting (O’Donoghue and Rabin, 2015), and salience theory (Bordalo et al., 2012) instead predict a sawtooth: effort rising steeply into each threshold, then resetting once it is crossed.

That is what we find. Pilot performance rises sharply above baseline from around 4–5 months before the quota is met. The additional output induced per pilot approaching the quota is on the order of six victories—substantial, given that the median pilot in our data scores just three in his entire career. Once the medal is received, most pilots scale back drastically, returning to their former rhythm. We find no evidence of increased pilot exits during the period of elevated effort.

Identifying these effects requires a novel procedure. Only pilots who score victories (a flow) can approach the quota (a stock), so performance spikes mechanically near any cumulative threshold—real or placebo. We therefore difference the performance path around the real quota against paths around placebo quotas, removing the mechanical effect and regression to the mean. A Monte Carlo experiment verifies that the procedure recovers true effects.

A complementary strategy exploits changes in the quota level. Quotas were adjusted by front over the war for reasons unrelated to any individual pilot’s performance; an increase could abruptly push pilots out of the “zone” of medal eligibility. Instrumenting “in zone” status with these changes yields results similar to our event-study procedure.

Section 3 turns to the organization’s problem: sustaining the value of its awards. The medal behaves like the fashion good in Pesendorfer (1995). Ranking medalists by ability—

permanent productivity purged of peer spillovers and environmental factors (Mas and Moretti, 2009)—we show that the ability rank of recipients declines over time. Each medalist thus watches his award “cheapen” as lower-ranked pilots receive it. The cheapening rationalizes the organization’s dynamic introduction of new medals—and ensures that each new award induces effort from high types.

Section 4 analyzes the resulting rat race. Four additional medals were introduced during the war, with strictly greater quota requirements, and awarded only in sequence. Pilots repeat the effort increase seen for the first medal quota as they approach the second, which we show by stacking event studies around the two quotas: the run-up to the second quota is at least as large as the first, and performance again falls back once it is met. “Making it”—once—is insufficient to nullify the desire for additional distinction.

Pilots respond even to the mere *introduction* of new medals. Using a generalized synthetic control approach (Xu, 2017), we show that medal-sensitive pilots—those who reacted strongly to quota proximity—increase their effort immediately after new medals are announced, relative to less sensitive pilots, even when far from qualifying themselves. Together, these results suggest that the creation and perpetuation of the status system was broadly effective for the Luftwaffe.

Section 5 probes the mechanism and computes aggregate magnitudes. Pilots with alternative sources of prestige—holders of unrelated medals, senior officers and noblemen—respond *less* to an approaching quota, in line with signaling models (see, e.g. Spence, 1973): strong existing signals weaken the incentive to invest in costly new ones. Medals thus act as substitutes for other status signals, underscoring status competition as the mechanism. In aggregate, the quota system caused 3,100–5,800 additional victories—6–11% of the total in our data.

To which organizations should these results travel? Our setting has three key features: pay cannot reward marginal performance, output is measurable and visible to peers, and the status hierarchy is credible. Militaries, civil services, and other bureaucracies share these features, as do academia and volunteer communities. But the mechanism does not require that money be unavailable. Microsoft has run its “Gold” and “Platinum” clubs since 2008, alongside its conventional bonus program (Sales & Marketing Management, 2014). The Gold Club honors more than two thousand employees each year, the Platinum Club roughly 150; both confer public recognition and exclusive experiences. Salesforce’s Trailhead platform

rewards training with publicly observable badges and ranks. For roughly eight years, Ranger was the top rank. In 2022, because Trailblazers “continued to learn and earn”, Salesforce added five more exclusive ranks, culminating in “All Star Ranger” (Salesforce, 2022). Firms with full access to financial instruments thus build purely symbolic status hierarchies of the kind we study—and extend them when their best people top out.

The advantage of our setting is to isolate the status mechanism itself: with no financial rewards attached to aerial victories, the responses we measure are driven by status alone. Nor are they well explained by wartime patriotism or ideology, which cannot account for outperformance concentrated precisely when it earns individual distinction, and weakest among pilots who already hold status from other sources. The features of our setting that do not travel—mortal stakes, a young and risk-tolerant workforce, and the absence of outside options—we discuss in the conclusion.

Our paper contributes to the literature on organizations’ use of economic (Kamenica, 2012) and social incentives (Ashraf and Bandiera, 2018) to address contracting problems. Building on work on monetary incentives and performance (Prendergast, 1999), a strand of research establishes a role for social recognition and signaling (Kosfeld and Neckermann, 2011; Ashraf et al., 2014; Ager et al., 2022); Akerlof and Kranton (2005) highlights identity as a source of motivation in military settings. Our focus is the dynamic problem of extracting repeated effort. Frey (2007) and Frey and Gallus (2017) study awards and highlight the empirical difficulty of estimating their effect on performance—a challenge we address by exploiting the quota criteria that govern the supply of medals.

In tournament theory, Lazear and Rosen (1981) show how non-linear rewards, such as promotions—or medals—induce effort and risk-taking among workers heterogeneous in ability when output is costly to observe. Besley and Ghatak (2008) model status incentives directly: awards can induce effort at zero marginal cost to the principal, in a static setting, even when effort is non-verifiable. We demonstrate empirically that such systems remain effective in a dynamic setting. Gill and Stone (2010) and Eisenkopf and Teyssier (2013) incorporate fairness concerns and loss aversion into tournament design.

Our central incentive is a status symbol, connecting to the economics of status. Hopkins and Kornienko (2004) show theoretically that status rewards induce overspending on status goods. Empirically, Bursztyn et al. (2018) identifies significant demand for a higher-tier

good with otherwise identical utility to a lower-status alternative; in [Imas and Madarász \(2024\)](#), the possibility of excluding others from a good increases its desirability. The Knight’s Cross medals are likewise a tiered system of signals whose supply was restricted over time. [Moldovanu et al. \(2007\)](#), [Auriol and Renault \(2008\)](#), and [Rayo \(2013\)](#) characterize how organizations can create and assign status—categories, symbols, and sold signals—to motivate workers, but none considers a repeatedly extended hierarchy of the kind we study. Our data are consistent with two key predictions of [Moldovanu et al. \(2007\)](#): despite the zero-sum nature of status competition, several coarse status categories are optimal under concave ability distributions, and the very top performers should not be pooled with lower types.

We also speak to the literature on dynamic incentives and cyclical effort. [Pesendorfer \(1995\)](#) explains how a status signaling game between consumers and a monopolist generates fashion cycles; we provide evidence that a similar process can guide a system of status-based worker incentives. [Misra and Nair \(2011\)](#) document output spikes just below quarterly sales quotas; we likewise find that workers do not smooth effort in response to discrete, quota-driven incentives. [Genakos and Pagliero \(2012\)](#) show that interim rank drives risk-taking among competitive weightlifters.

Finally, our rat race relates to signaling, goal-setting, and hedonic adaptation. [Akerlof \(1976\)](#) formalized the rat race: working harder signals a superior type that cannot otherwise be observed. [Landers et al. \(1996\)](#) document the resulting overwork in US law firms, but in a one-off promotion without the dynamics we study. Research on hedonic adaptation documents habituation following major life events ([Brickman and Campbell, 1971](#); [Frederick and Loewenstein, 1999](#); [Bottan and Perez Truglia, 2011](#); [Perez-Truglia, 2012](#)). [Genicot and Ray \(2020\)](#) summarize the role of goal-setting in guiding effort; [Allen et al. \(2017\)](#) document bunching at reference marathon times that carry no external incentives. In [Bénabou and Tirole \(2011\)](#), accumulated status leads to a “self-defeating” over-acquisition of yet more status—a treadmill that our expanding range of status symbols amplifies.

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

This section briefly reviews the role of the German air force (*Luftwaffe*) in World War II, describes the system of Knight’s Cross medals, and summarizes 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 built a new one under the energetic leadership of Hermann Göring, a former fighter ace and squadron peer of Manfred von Richthofen. The Luftwaffe was designed mainly to provide close air support to the army, which required air superiority over the battlefield and thus a focus on fighter and ground attack aircraft. Germany invested heavily in advanced aircraft, such as the Bf 109 fighter and the Stuka dive-bomber, and developed new tactics, such as the four-finger formation—both tested and refined during the Legion Condor’s support of General Franco in the Spanish Civil War.

The campaigns in Poland and France quickly showed the merits of close air support and air supremacy. In the Battle of Britain, however, the Luftwaffe met its match in men and materiel, failing to subdue the RAF ([Bungay, 2010](#)). In the opening phase of the campaign in Russia, it 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, though in the early years of the campaign German pilots and machines held a clear qualitative edge. On the Western front and over the Reich itself, air superiority was lost by 1944, when long-range fighters escorted Allied bombers. After the spring of 1944, the Luftwaffe was a shadow of its former self: Allied bombing destroyed fuel plants, training suffered, and pilots were dying at a rate of 25% or more per month ([Murray, 1983](#)).

The campaigns in France and Britain saw a small, select group of pilots attain *Experten* (ace) status, having shot down 15 or more enemy planes. The one-sided air war in Russia made total Luftwaffe victory claims surge after 1941. Because German pilots were never rotated out, even after extended tours (“fly till you die”), top German tallies exceed those of Allied pilots by a large margin. Detailed comparisons accounting for time in combat and

enemy numbers nonetheless fail to show any major advantage of the best German over the best Allied pilots.⁴

Victories mattered greatly for a pilot’s status among his peers, in the air force, and in society at large.⁵ The High Command dedicated substantial resources to tracking and verifying individual victory claims, using gun camera footage when available. Every victory was entered in the pilot’s logbook—the enemy plane’s type, how it was shot down, whether the impact was observed, and the grid coordinates of the wreckage—and had to be witnessed by a fellow pilot or confirmed by a 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](#)).⁶

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. The original Knight’s Cross of the Iron Cross, introduced in September 1939 at the outbreak of the war, was an all-purpose medal awarded with particular emphasis on individual bravery ([Williamson and Bujeiro, 2004](#), p. 3).⁷ It topped a “ladder” of awards—only recipients of the Iron Cross 1st and 2nd class qualified. The Iron Cross itself, introduced during the War of Liberation in the early 19th century, was deliberately conceived as a meritocratic award, bestowed without regard for rank or noble origin.

As the war continued, numerous high performers—successful U-boat captains, tank “aces”—received the KCR, and the armed forces, with the Luftwaffe particularly vocal, lobbied for additional medals to recognize continued performance. Higher-ranked variants followed: the Knight’s Cross with Oak Leaves in 1940, with Oak Leaves and Swords and with Oak Leaves, Swords, and Diamonds in 1941, and with Golden Oak Leaves, Swords, and Diamonds in

⁴[Simkin and Roychowdhury \(2006\)](#) examine this question for German pilots during WWI, when similar incentives obtained, and find no sign of outperformance.

⁵Propaganda regularly highlighted the accomplishments of German pilots. Pilots’ pictures were distributed as postcards, similar to portrait pictures of movie stars.

⁶Total German victory claims are typically less than the sum of Allied aircraft destroyed, according to Allied records, which suggests underclaiming.

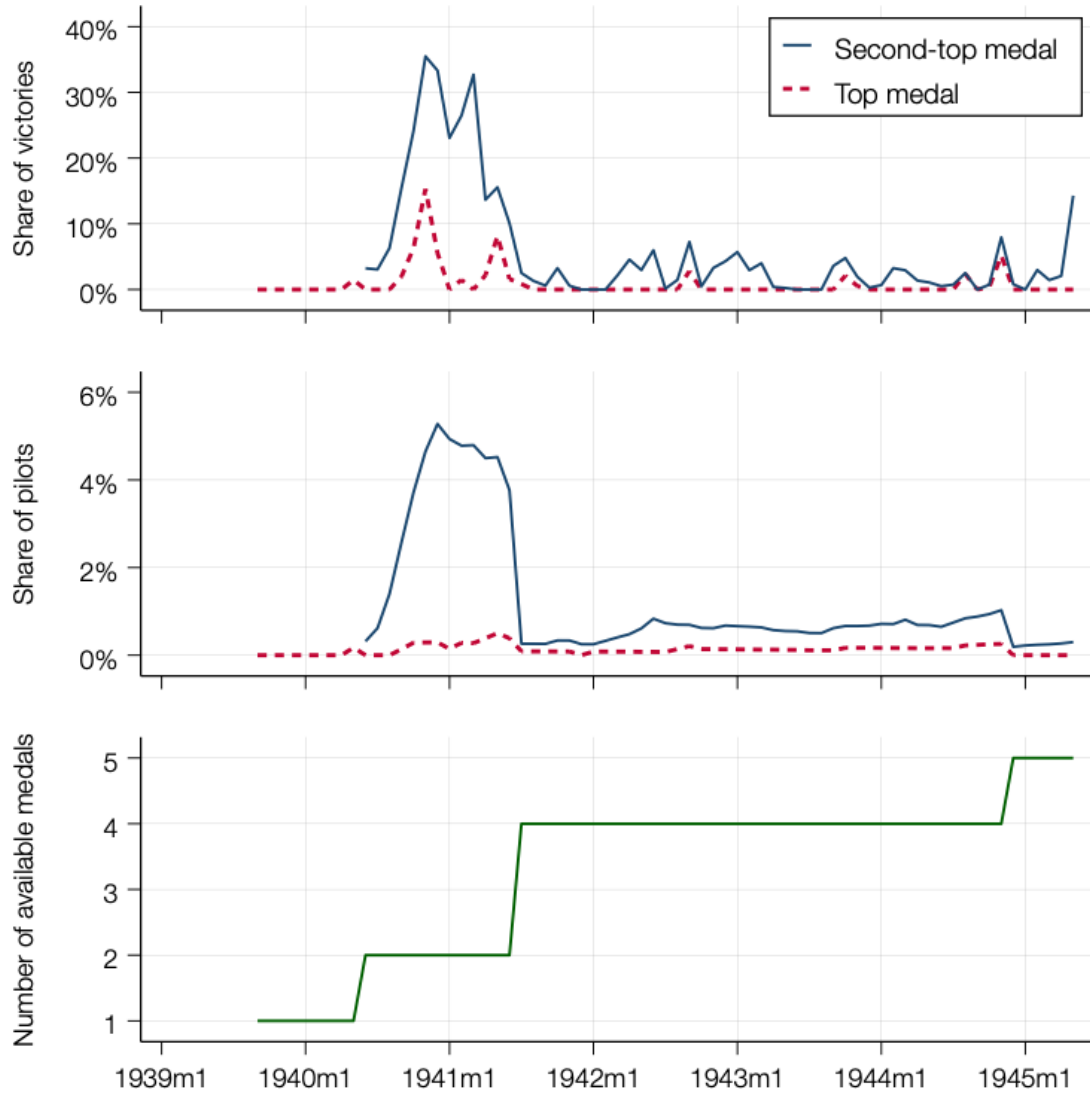
⁷While awards to high-ranking officers occurred, the emphasis on personal bravery increased sharply during the war. From 1942 onwards, less than half of the medals awarded went to officers, compared to around four-fifths in 1940.

1944.

Knight's Cross medals were worn around the neck, a prominent part of a soldier's uniform; the medal itself bore no indication of when—or under which quota—it had been earned. Remarkably, the medals partly reversed the military hierarchy of deference: even high-ranking officers often saluted decorated NCOs. Only 7,175 were ever awarded, including 882 Oak Leaves, 159 Swords, 27 Diamonds, and one Golden Oak Leaves. Most awards went to the army, but the air force had a high concentration of awardees relative to eligible personnel. A total of 558 Knight's Cross-level awards went to fighter pilots.

Figure 1 provides an overview of the Knight's Cross system. The lower panel shows the staggered rollout of additional medals; the middle panel, the share of pilots holding the top and second-top medals over time. Close to one in twenty pilots gets close to “topping out” of the incentive ladder before it is extended, and this rising concentration of medalists preceded each new introduction. The upper panel shows that medal holders account for a share of monthly victories highly disproportionate to their share of pilots—controlling their incentives was therefore crucial. We henceforth refer to the medals as medal 1, medal 2, etc., in ascending order of rank.

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, 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 for nomination. These targets were informal—not

a strict “quota”—but guidance was “brought to the attention of field commanders”, and every nomination involved discretion, passed up and endorsed by each level of the chain of command before final acceptance (and often personal award) by Hitler himself. Pilots were keenly aware of the targets. Their aircraft’s tails displayed their “tally”, keeping everyone aware of their own and their peers’ standing. [Obermaier \(1966\)](#) argues that the quota system was explicitly created as “psychological bait”.

The initial quota for medal 1 was 20 victories, a threshold reached by just six pilots in the first twelve months of the war. As the war went on, the quota was adjusted to preserve the prestige desired by the High Command: a separate, higher quota was introduced on the Eastern front, where victories came easily against inexperienced and poorly equipped Soviet pilots, and quotas were lowered as the war turned against Germany. Medals 2 and 3 had their own quota requirements. We take the specific levels from the historical account in [Obermaier \(1966, pp. 21–23\)](#); [Figure A1](#) shows the substantial variation over time, reflecting the High Command’s desire for a consistent standard amid changing combat conditions. The quota for medal 3 at one point rose to 250 victories—an astounding target, given that the median pilot in our dataset (already restricted to pilots with at least one victory) finished the war with just three.

The medal system could be modified as priorities changed. From 1943, Allied bombing raids became a growing threat: bombers, especially the USAAF B-17 “Flying Fortresses”—equipped with up to 13 Browning 0.50 caliber machine guns and increasingly accompanied by fighter escorts—were far harder to shoot down than single-engine fighters and accounted for a high share of German aircraft losses. To align incentives, a points system (*Punktwertung*) was introduced that rewarded pilots according to the type of adversary and their role in its destruction. Two reasons make us skeptical of its importance relative to the quota system. First, [Obermaier \(1966, p. 21\)](#) highlights pilots’ resistance to the new system: they disliked the distraction from the purity of fighter versus fighter combat and the complications of crediting the group effort needed to attack bomber formations.⁸ Second, the predictive power of the victories quota for medal awards shows little decline after the points system’s *de jure* introduction in October 1943. We use the victories quota as the salient target for pilots throughout the analysis.

⁸The air force tried to address some of these issues by awarding points for forcing bombers out of formation, in addition to destroying them.

The dispatch mentions studied in [Ager et al. \(2022\)](#) and the Knight’s Cross medals are complementary incentives, and the two are only weakly correlated ($\rho = 0.1567$). Three differences make medals particularly useful for our purposes. First, medals were permanent and visible—worn prominently around the neck on all occasions, including while flying ([Williamson and Bujeiro, 2004](#))—in contrast to the fleeting recognition of a mention. [Figure A2](#) shows medal 1. Second, the award was not only highly prestigious—underscored by the many awards Hitler made personally—but also predictable and “attainable”: the informal quota system made it fairer and more transparent than mentions, and, among fighter pilots, approximately 6.6 times more Knight’s Cross medals were awarded than mentions in dispatches.⁹ Finally, the medals form an explicit hierarchy, extended throughout the war, in which each award was available only to holders of the preceding one—exactly the structure needed to study a rat race in pilot performance.

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 panel from several comprehensive secondary sources based on microfilm data from the German Federal Archives. The data covers daytime fighter pilots (engaged in air-to-air combat) from September 1939 to May 1945, limited to pilots who score at least one aerial victory. Summary statistics for the pilot–month panel ($N = 82,348$) are in Panel A of [Table A1](#), with pilot-level characteristics in Panel B ($N \text{ pilots} = 5,081$).

Our primary outcome is the monthly victory rate: counts of enemy aircraft shot down and verified by High Command. We also study exit, a dummy coding departure from the panel before the end of the war. Around half of the exits are documented in two separate sources; two-thirds represent the pilot’s death, the remainder his being shot down or wounded seriously enough to end his flying.

⁹Mentions 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 7,175 Knight’s Cross medals were ever awarded. For fighter pilots in particular, our data has 85 mentions and 558 Knight’s Cross medals.

For a subsample of pilots, we observe flying activity directly: [Ager et al. \(2022\)](#) collect the flight logs of 71 pilots from private collections, recording each flight flown and whether it was a combat sortie. Surviving logbooks naturally skew toward successful pilots—precisely the group the medal system targeted. We use this subsample to corroborate our interpretation of the victory data.

For our mechanism results, we construct a measure of pilot ability, following [Mas and Moretti \(2009\)](#) to recover individual permanent productivity in the presence of peer spillovers. To handle the large fraction of unique peer combinations relative to observations, we apply the network reduction procedure of [Hamilton et al. \(2022\)](#), which recovers time-invariant ability for most pilots; Appendix B provides details. We use “ability” and “quality” interchangeably throughout. Our treatment-effect estimates rely on within-pilot variation only; the ability measure serves to classify pilots and to study mechanisms.

Medals

We collect the identity of medalists and the timing of Knight’s Cross awards from *Lexikon der Wehrmacht* (<https://www.lexikon-der-wehrmacht.de/>; last accessed August 2026), a website housing extensive databases on German military history, which records each medalist’s full name, branch, unit, and rank, along with the award date and medal type. We manually harmonize the spellings of names, ranks, and units and merge the medal database into our panel. We verify the data against an alternative source, the Kracker Luftwaffe Archive (<https://aircrewremembered.com/KrackerDatabase/>; last accessed August 2026), finding the two highly similar but preferring the *Lexikon* for its Knight’s Cross coverage; from Kracker, we collect the lower-tier *Deutsches Kreuz* medal. Panel B of [Table A1](#) shows that around 8% of pilots achieved medal 1, 2% medal 2, and 1% medal 3. Only 6 daytime fighter pilots received medal 4, and none received medal 5, the highest medal,¹⁰ so these medals are too rare 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 are needed to complete a month-by-front-by-medal panel of quota levels. Our preferred interpretation was hand-coded from the German text by a research assistant

¹⁰The highest medal was only awarded to a ground-attack pilot, Hans-Ulrich Rudel.

without knowledge of our research design; three other coders worked from the source’s parallel English translation. Discrepancies across coders are minor and have no substantive impact on our analyses. We show robustness to an alternative coding of the quota level in the Appendix.

2 Medal Chasing

This section examines whether pilots “tried harder” and had greater success as they approached the quota for the first medal, the Knight’s Cross. We describe the non-standard econometric problem; propose an estimator that corrects mechanical bias using placebo quotas, validated in Monte Carlo simulations; collect the argument for interpreting our estimand as effort; and present the bias-corrected estimates.

2.1 Empirical design

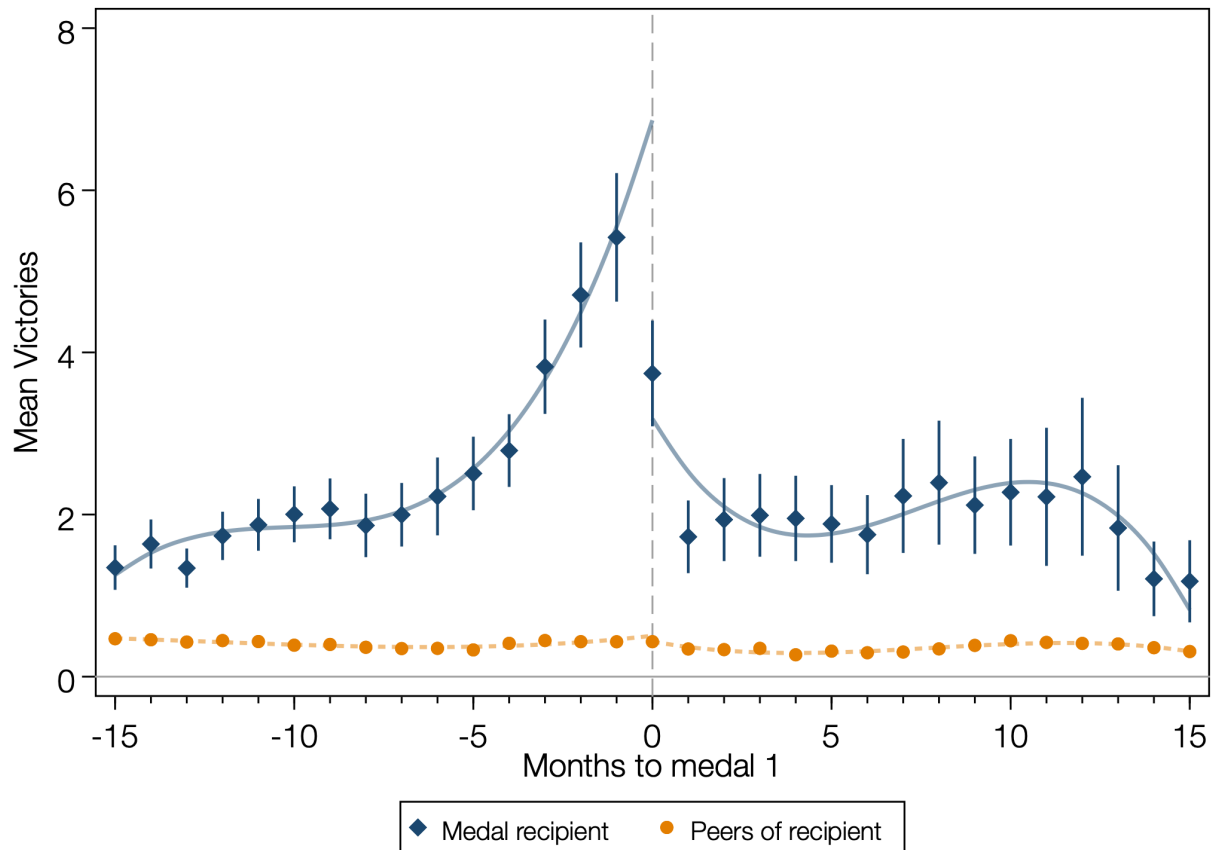
Do pilots respond to the prospect of receiving the first medal in the Knight’s Cross hierarchy? In [Figure 2](#), we plot the mean victories of pilots about to receive the Knight’s Cross (in blue) against peers at the same airbase and calendar time (in orange). The raw data show a suggestive spike prior to medal receipt and a sharp decline after—but this exercise conditions on the outcome (medal receipt).

To measure how much of the pre-medal increase in performance reflects additional effort, we proceed in two steps. First, we define event time by the medal *quotas* rather than actual receipt: pilots did not know if or when they would receive the medal, but they knew their distance to the relevant quota. Second, we purge the upward bias that arises because scoring victories mechanically brings a pilot closer to the target.

The quota strongly predicts when medals are awarded. [Figure A3](#) plots the likelihood of receiving medal 1 against distance to quota (cumulative victories minus the quota level). The likelihood of an award (“medal risk”) peaks just after pilots reach the quota, consistent with flexible yet relevant eligibility criteria: the quotas’ imperfect bindingness was an explicit part of the system, and the slight offset of peak risk from zero likely reflects procedural delays

as nominations passed up the chain of command (Figure A4). In the language of Frey and Gallus (2017), Knight’s Cross medals are confirmatory rather than discretionary awards. We therefore consider event time relative to meeting the quota—the timeline pilots could anticipate and control.

Figure 2: Raw victory rates over medal event time



Notes: Figure shows mean monthly victories in blue over time relative to winning 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.

We hypothesize a performance profile with two key features. First, the quota—a target for the *stock* of victories—lets pilots anticipate the likelihood of receiving a medal, so they should increase their *flow* of victories as they approach it. We interpret the sum of the quota-timed effects up to and including the month the quota is met as an “in zone” effect, measuring how strongly the medal motivates pilots. Second, performance should fall sharply once the quota is exceeded, since the medal is received or imminent—confirming that the profile is

tied to the medal rather than spurious. Consider the population-level empirical model:

$$\bar{Y}_j = \theta_j + b_j \quad (1)$$

Define the event time E_i as the period t in which pilot i meets the medal quota.¹¹ $\bar{Y}_j$ is the mean victory rate of pilots at event time $j = t - E_i$ (so that $j < 0$ before the quota is met), θ_j is the quota-timed performance effect—our effort estimand, an interpretation we defend in Section 2.2—and b_j is the bias term.

The econometric challenge in recovering θ_j is that we expect non-zero b_j : systematic positive errors from two distinct selection effects. The first is *mechanical* selection: positive shocks advance the event-time index, since each victory today brings the quota closer—informally, the challenge of regressing a flow on a stock. Even under the null hypothesis $\forall_j : \theta_j = 0$, $\bar{Y}_j$ is large just before the quota because only positive shocks move a pilot toward it; at a minimum, the crossing month must contain at least one victory. This relates to Ashenfelter’s Dip, in which low earnings mechanically select workers into training programs (Ashenfelter, 1978), though our problem is distinct: we estimate pre-treatment effects, so the design is fundamentally not difference-in-differences.

The second is a *survivorship* effect: latent pilot ability drives both the victory rate and survival. Pilots who meet the quota are positively selected for ability, as the persistent level gap between the blue and orange lines in Figure 2 shows. Fixed effects absorb this level difference, but further correction is needed for compositional change as event time advances, because higher-ability pilots survive longer.

We posit a linear panel model, adjusting for the nuisance b_j terms—the “mechanical effect”. The standard regression equation is:

$$\begin{aligned} Y_{i,t} = & \alpha_i + \delta_t + \lambda_{i,t} + \varphi_{i,t} \\ & + \sum_{j \in \mathcal{J}} \mu_j \mathbb{1}\{t - E_i = j\} \\ & + \rho Y_{i,t-1} + \varepsilon_{i,t} \end{aligned} \quad (2)$$

Here α_i and δ_t are unit and time fixed effects, and the indicators $\lambda_{i,t}$ and $\varphi_{i,t}$ adjust for

¹¹Medal 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.

combat front and for holding the medal. 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 consider a Poisson count model as a robustness exercise. The residual error is $\varepsilon_{i,t}$.

We fully saturate (2) with treatment lags and leads, with a leave-out period of $j = -15$ —chosen to precede anticipation effects—and lags and leads $|j| > 15$ accumulated into end bins. The μ_j terms nonetheless remain contaminated with positive bias from the mechanical effect.

Our solution exploits the fact that the mechanical effect can be recovered from quotas that carry no treatment effect by construction: placebo quotas.¹² As a pilot approaches a placebo quota, his effort is unchanged, but the mechanical effect appears in an equation like (2); differencing it out of the overall event-study effects recovers the clean effect. Specifically, we generate multiple placebo quotas¹³ and stack the true lags and leads ($p = 0$) with the placebo lags and leads ($p \in \{1, \dots, P\}$). The clean effect θ_j is then estimated by interacting the event-study indicators with an indicator for the true event time:

$$\begin{aligned} Y_{i,t,p} = & \alpha_i + \delta_t + \lambda_{i,t} + \varphi_{i,t} \\ & + \sum_{j \in \mathcal{J}} \left[b_j \mathbb{1}\{t - E_i^p = j\} + \theta_j \mathbb{1}\{t - E_i^p = j\} \cdot \mathbb{1}\{p = 0\} \right] \\ & + \rho Y_{i,t-1,p} + \varepsilon_{i,t,p} \end{aligned} \quad (3)$$

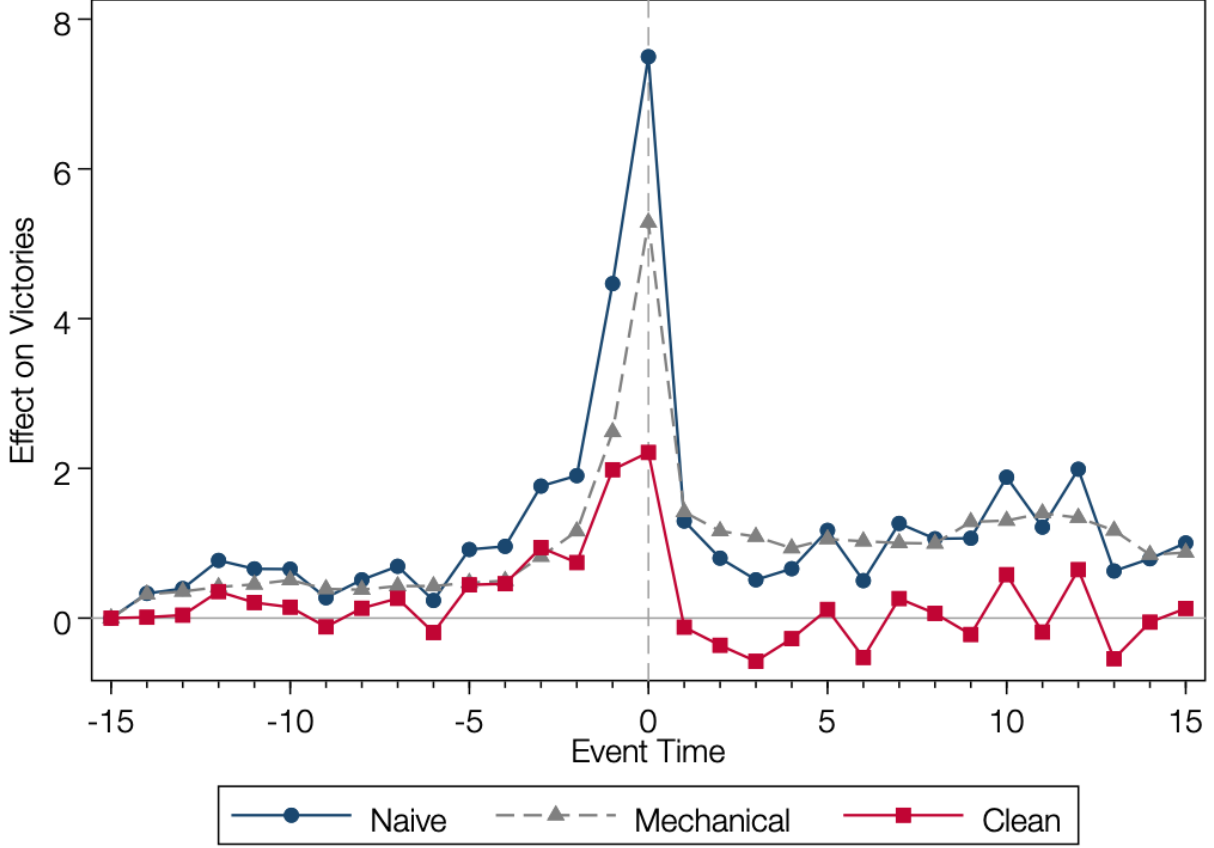
Note that the data is now stacked at the (i, t, p) level, with the restriction that $Y_{i,t,p} \equiv Y_{i,t}$ is repeated across p . This stacking mechanically correlates $\varepsilon_{i,t,p}$ within (i, t) . As such, for inference, we cluster at the pilot or (coarser) squadron level.

Figure 3 separates the mechanical effect b_j from the “clean” causal effects θ_j , estimated with 25 placebo quotas per pilot. The naive estimates correspond to the μ_j in (2) over the original (non-stacked) data; imposing $\mu_j = \theta_j + b_j$, the mechanical effect is the residual difference. As expected, it is everywhere positive and peaks strongly just before the quota.

¹²This is in the spirit of Borusyak and Hull (2023), who “recenter” estimates after specifying a model of nonrandom exposure to shocks.

¹³We draw placebo quotas uniformly from $q_{it} \cdot [\underline{\mu}, \bar{\mu}]$, where q_{it} is the real quota facing a given pilot. We set $\underline{\mu} = 0.1$ and $\bar{\mu} = 2.0$.

Figure 3: Decomposition of the event study into mechanical and clean effects

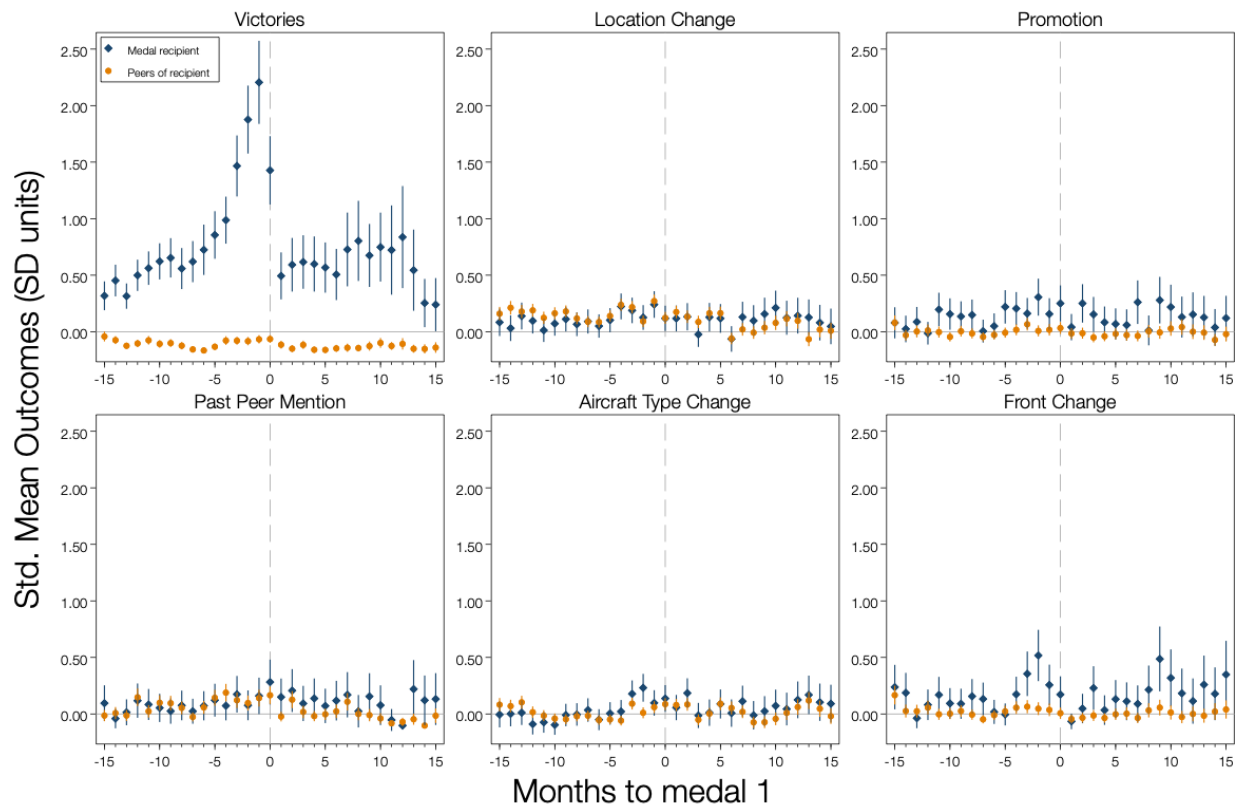


Notes: Figure shows the decomposition of coefficients from an event study over time to meeting the quota for medal 1 under our mechanical-correction procedure. The *Naive* series plots the coefficients from a “naive” estimation of (2) on 82,348 observations without any adjustment. The *Clean* series 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 bias term b_j .

Our estimate $\hat{\theta}_j$ from (3) is consistent for θ_j in (1) under three identification assumptions. First, the leave-out period, $j = -15$, is sufficiently distant from the quota that pilots have not yet adjusted their effort; a baseline level of effort elevated by the medal system’s existence is partialled out by pilot fixed effects. Second, averaged across draws, the placebo quotas share a common error structure with the real quotas: $E[b_j^{\text{placebo}}] = b_j^{\text{real}} \quad \forall j \in \mathcal{J}$. We support this by drawing placebo quotas from a distribution empirically similar to the real quotas—defined on the same cumulative victories process with comparable support, they induce the same flow-on-stock conditioning and survivorship selection. We do not prove the equality formally, but verify by simulation that the procedure recovers true effects under a data-generating

process with these features (Appendix C). Third, other determinants of performance (e.g., tactical reassignments, weather, materiel) are not systematically timed relative to the quota. Beyond the conditional exogeneity provided by rich fixed effects, Figure 4 shows directly that the rates at which pilots change airbase, are promoted, are treated by the past peer mention from Ager et al. (2022), change aircraft type, and change combat front do not vary over event time around medal 1 and do not deviate from squadron-peer trends—in sharp contrast to the divergence in victories.

Figure 4: 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 medal 1. The orange line shows the mean value for peer pilots who fly in the same squadron and month as the focal pilot. Before averaging, each variable is standardized (z-scored) using its mean and standard deviation in the estimation sample, so that all panels share a common y -axis measured in standard-deviation units. The variables (left-right, top-bottom) are the victory rate, and indicators for being assigned to a new 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.

A Monte Carlo experiment verifies that this approach recovers the effort effects θ_j . Our DGP broadly matches the size and empirical features of the pilot-month panel and includes

endogenous exit driven by risk-taking and randomness; the full specification is in Appendix C. We generate a logistic-shaped medal anticipation effect under the alternative hypothesis of a positive in zone effect, and also test the estimator under the null $\forall_j : \theta_j = 0$. Figure A5 shows the results of 50 simulations, each with $P = 25$ placebo quotas per pilot. The raw estimator (2) overshoots the true pre-quota effort by a factor of 3 and exhibits a spurious positive effect *after* the quota is reached. By contrast, the clean estimates track the true effect closely, with confidence bands covering θ_j across the event window under both hypotheses. We apply the mechanical-correction procedure to all individual-level results that follow.¹⁴

2.2 Effort estimand

Before presenting estimates, we collect the argument for interpreting θ_j as effort. Effort is never directly measured; monthly victories reflect ability, combat conditions, luck, and effort. We make no claim to decompose the *level* of output. Our claim is weaker: the quota-timed component of within-pilot performance reflects effort. For this interpretation to fail, a non-effort component must vary systematically with time to the quota. We consider each in turn.

Ability is absorbed by pilot fixed effects: the treatment effects on victories are estimated from within-pilot variation. A richer measure of ability, net of additive peer spillovers (Mas and Moretti, 2009) is used to classify pilots in the mechanism analysis of Section 3.

Combat conditions are absorbed by calendar-time fixed effects and front controls. The residual concern is that conditions improve precisely as a pilot approaches the quota—say, because commanders steer promising missions his way. Figure 4 shows directly that reassignments, promotions, aircraft changes, and front changes are all flat over event time, in sharp contrast to victories.

Luck is differenced out by the mechanical-correction procedure. Lucky streaks carry a pilot toward any cumulative threshold, real or placebo, so the placebo quotas reproduce selection on luck—including “hot hands” and the subsequent regression to the mean. We do not require luck to be classical; we require only that it does not distinguish real thresholds from

¹⁴Note that this correction results in regressions with approximately $(P + 1) \cdot N = 26 \cdot 82,348$ observations. Our coverage of θ_j is complete except in the crossing month under the alternative hypothesis, where the adjusted estimate recovers about half of the true effect.

placebo ones.

Joint production with peers could shift credit toward a pilot nearing the quota without any change in his behavior. Squadron peers of near-quota pilots show no decline in victories (Figure 4)—no one “hangs back”—and no increase in exit risk (Figure A6)—no one absorbs risk on the ace’s behalf. We conduct a systematic test for complementarity between squadron members using a Shapley decomposition of squadron output. Appendix D reports our findings of no complementarity in performance: the complementarity term is small and statistically insignificant, and its 95% confidence interval rules out that complementarity accounts for more than 8.7% of all victories.

The scope for effort discretion is visible in the raw data. Among pilots observed within five victories of the quota—just above or just below—monthly victories are 3.1 to 4.6 times higher for those who have yet to receive the medal than for those who already hold it (Figure A7). At the same point in the tally, output differs massively with whether the incentive is still live: pilots can evidently modulate their performance.

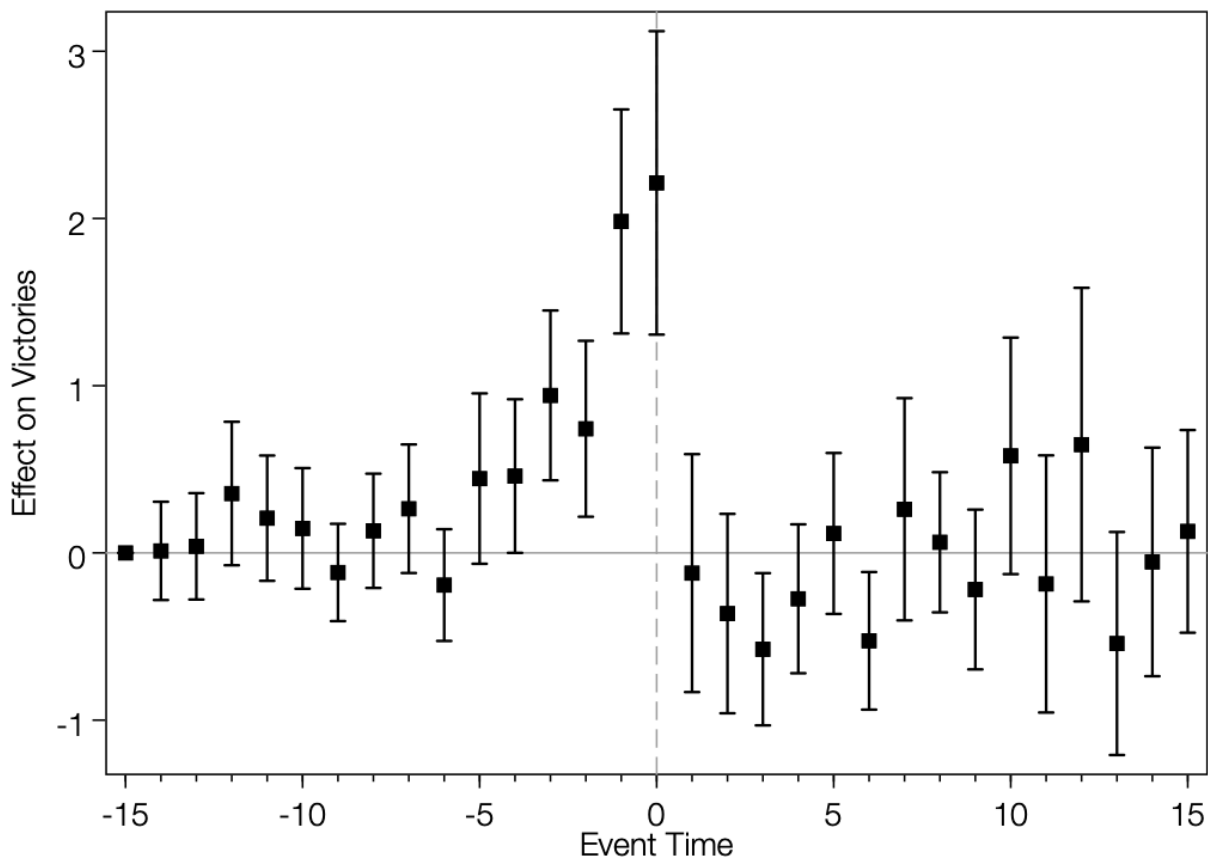
Finally, we observe behavior directly for the flight-log subsample (Section 1). Sorties flown measure the extensive margin of effort; victories per combat sortie, its intensity. Residualized victories load strongly on monthly sortie counts (Figure A8), validating that the variation we study tracks a physical input rather than noise. Figure A9 plots both margins against distance to the quota at the start of each week: victories per sortie roughly double over the final twenty victories, and sorties flown, relative to other active pilots, peak as the quota is crossed—while the medal decision hangs in the balance. Pilots work more, and work harder per flight, exactly when the medal is at stake. These profiles carry the same mechanical selection as the raw event study, so we read them as corroboration rather than causal estimates.

2.3 Individual performance results

The event study in Figure 5 plots the $\hat{\theta}_j$ terms and 95% confidence intervals from the mechanical-correction procedure, clustered by pilot; we interpret the coefficients as effort effects, as argued in Section 2.2. Performance accelerates 4–5 months before pilots meet the quota, peaking at around one standard deviation of monthly performance. The total induced

effort—the area under the treatment effect curve—is 6.34 (s.e. = 0.63). The median pilot scores just three victories in his entire career, so each medal is equivalent to deploying two additional pilots for the duration of their service (at vastly increased cost in equipment and gasoline); note that the result is driven by pilots of much higher “quality” than the median. Immediately upon receiving the medal, effort drops sharply—important because it shows that “hot streaks” of favorable combat conditions do not drive our results. Pilots exercise calculated precision in their exertions in response to the medal quota: effort also falls sharply for pilots who receive the actual medal while still below the quota (Figure A7).

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



Notes: Estimated average treatment effects on victories per month as a function of their time relative to meeting the medal 1 quota. The event study model (3) includes 15 lags and leads of treatment, with the 15th lead ($j = -15$) as the reference period; periods further from the quota are absorbed into separate end-bin indicators. 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. Fixed effects are pilot and period. Standard errors clustered at the level of fighter pilots with 95% confidence intervals indicated by the capped spikes.

We interpret the rapid pace of victories just before the quota as increased effort: entering “the zone”, each additional victory carries the promise of a highly coveted medal. That out-performance concentrates precisely where it earns individual glory further suggests that patriotic or ideological factors are not the main motivating force in our sample.

The return to baseline rhythm after achieving the quota—thereby earning the medal—can reflect two effects. First, the immediate incentive is removed, and the next quota is typically far off. Second, medalists were sometimes deployed on speaking tours and other propaganda duties upon receipt. Since we cannot systematically distinguish such reassignments from effort withdrawal, we do not overinterpret these post-treatment effects.

Robustness

We summarize several robustness tests in [Figure A10](#), alongside the baseline specification described above.

A parsimonious specification of (2)—without covariates for medal status or the first lag of the outcome—rules 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 $j = 1$.

Ace pilots yield the same result: we re-estimate the effect only on pilots in the top skill quintile of the [Mas and Moretti \(2009\)](#) measure (795 pilots). The effort interpretation is easier in this group, since [Figure 6](#) (discussed below) shows that aces can calibrate their effort and risk-taking in a way that lower-skill pilots cannot. Differences from the full-sample results are minor, given that aces dominate the treatment group.

The quota coding does not drive the result: an alternative coding by a different reader of the source text ([Obermaier, 1966](#)) adopted, unprompted, a more conservative interpretation of the number of quota level changes. The two codings correlate at $\rho = 0.9559$ at the front $\times$ month level and show only minor differences in the event study.

The quota’s informality could inflate anticipation effects: in exceptional cases, pilots just below the quota received the medal early for significant accomplishments—and accomplish-

ments preceding a discretionary award would be unsurprising. We address this with an alternative specification in which the event requires both meeting the quota *and* holding the medal, cleaning the pre-period effects of the exceptional awardees so that treatment reflects only the “standard” case. Under this definition, we cannot reject equality with the baseline except in period $j = 0$, where the effect is negative: affirmed by the medal, pilots can begin to “slack off” earlier.

Alternative estimation using instrumental variables

A distinct approach uses instrumental variables, exploiting that the quota level varies over time and is orthogonal to the error terms $\varepsilon_{i,t}$ for *individual* pilots. The High Command adjusted quotas—generally upward, as more pilots won medals—for entire theaters at a time, to preserve the medal’s exclusivity given combat conditions; the changes are thus independent of any one pilot’s performance, conditional on front and period fixed effects (Figure A1 shows the extensive program of changes). In Appendix E, we recover a static version of the in zone effect by instrumenting a pilot’s quota proximity with front-wide quota increases. The coefficients in Table A2 point in the same direction and are broadly comparable (though larger) in magnitude relative to the event-study estimates. The static approach is nonetheless less well-suited to identifying dynamic policy effects and requires additional assumptions.

Effects on risk taking and exit

Despite generally high baseline risk, we find no differentially higher pilot exit while in zone of the medal quota.

We estimate Cox proportional hazards models with exit—death, serious injury, MIA/POW status, or other severe incapacitation—as the failure event. Table 1 presents hazard ratios for the in zone treatment, which indicates that the pilot has come within fifteen victories of the quota for his next medal and has not yet received that medal. A pilot who meets the quota remains in zone until the award. Across stratum specifications for experience, front, fighter squadron, and medal status, and controlling for pilot quality, the estimates are statistically indistinguishable from unity; the 95% confidence intervals lie within 0.68 to 1.25. Figure A6

shows the raw data:¹⁵ exit rates are uncorrelated with quota distance, in stark contrast to the clear patterns for medal risk and victories, and the same holds for squadron peers in the same month—speaking against joint production or risk-sharing driving the results.

Being in zone thus raises performance without significantly raising exit risk. Pilots who reach the zone are already of exceptional ability, with the leeway to increase effort without excessive risk—precisely why the organization prioritizes motivating these aces in the medal system.

Table 1: Estimates of the in zone effect on exit

	Exit (0/1)				
	(1)	(2)	(3)	(4)	(5)
Inzone	1.068 (0.086)	1.015 (0.083)	0.938 (0.079)	0.881 (0.119)	0.899 (0.130)
Pilot quality		1.031* (0.016)	1.029* (0.016)	1.014 (0.021)	0.979 (0.023)
N	76,599	76,599	76,599	76,599	76,599
χ^2	0.66	4.00	3.63	1.13	1.66
<i>Strata</i>	Experience	Experience	Experience Front	Experience Front Squadron	Experience Front Squadron Medals

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The table presents hazard ratio estimates from Cox regressions. The failure event is pilot exit. Inzone indicates that the pilot’s cumulative victories are no more than 15 below the quota for his next medal, or at or above it, and that he has not yet been awarded that medal. Strata are defined based on experience, combat front, fighter squadron ID, and the number of medals held. The baseline hazard rate is allowed to vary flexibly across strata. All columns are estimated on a common sample of pilot-months with non-missing pilot ability. Standard errors (in parentheses) are clustered at the level of fighter squadrons.

¹⁵Note that the absence of such an effect is arguably unsurprising. [Ager et al. \(2022\)](#) showed that the most skilled pilots did not see an uptick in exit rates following a peer’s recognition, while effort increased. The opposite was true for lower-quality pilots. Arguably, the treated population in the medal setting is more similar to the highly skilled pilots in [Ager et al. \(2022\)](#).

3 The Organization’s Problem

Having documented the increase in effort around medal quotas, we turn to the organization’s problem: eliciting effort repeatedly through symbolic awards, despite the risks this entailed.

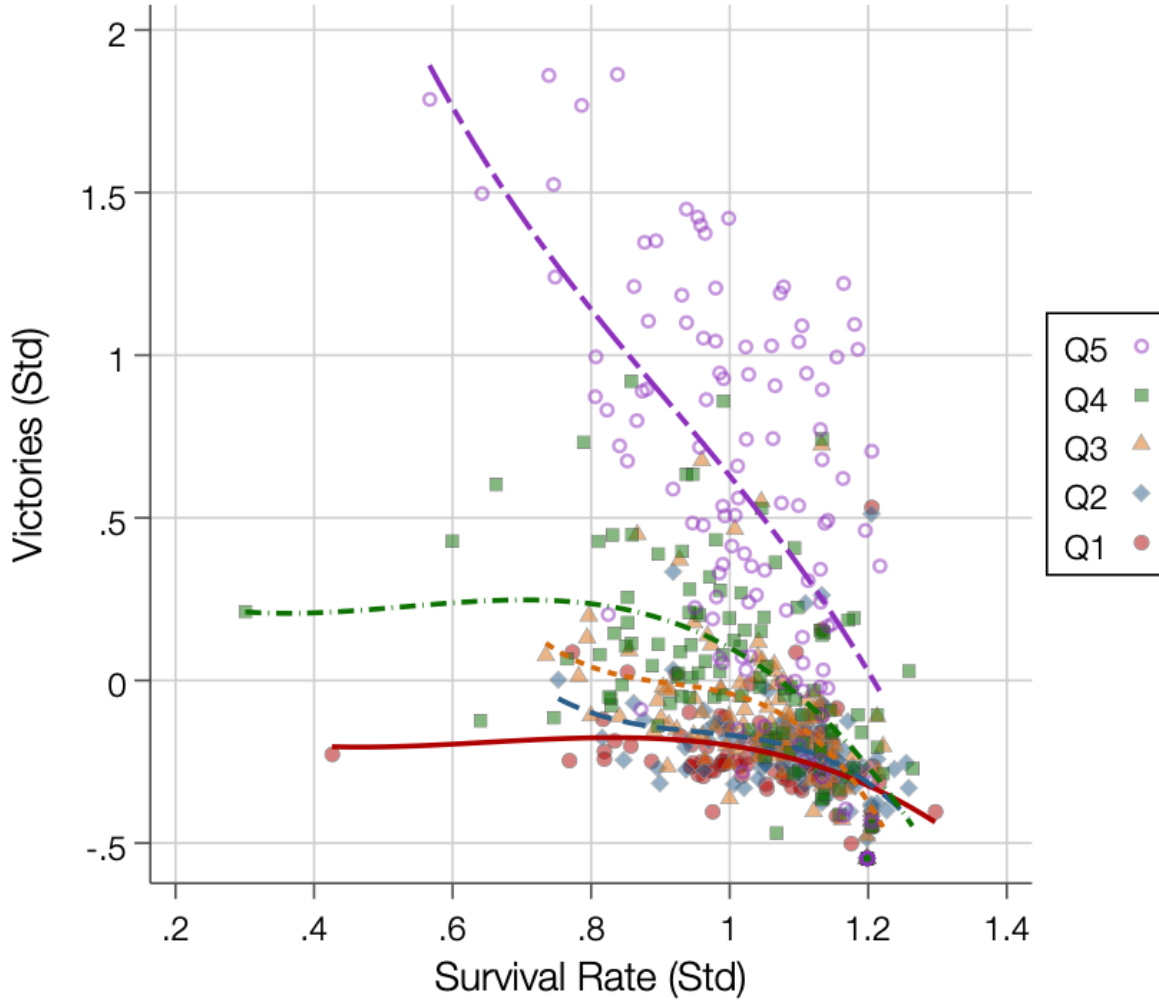
3.1 The cost of effort

Top pilots contributed a large share of the overall victory tally—motivating this group was important. Of the total 54,835 aerial victories in our data, 40.1% were scored by the 10% most able pilots by the [Mas and Moretti \(2009\)](#) measure; the top 1% alone accounted for 8.0%.

At the same time, all pilots—elite pilots included—had to take substantial risks to win in aerial combat. [Figure 6](#) illustrates the tradeoff as an empirical production possibilities frontier, aggregating the data to the month $\times$ front $\times$ ability quintile level. The x -axis plots the (standardized) average survival rate; the y -axis, the (standardized) average victory rate. The further left an observation lies, the greater the risks taken; the higher, the more enemy planes destroyed.

The organization wants pilots to perform as far to the top-right as possible in this space. Top-quintile pilots (lilac-dashed line) negotiate the trade-off best: every unit of risk yields many more victories. Second-top quintile pilots are already less than half as productive; for the third-top quintile, output per unit of risk again halves; bottom-quintile pilots show only a slight increase in victories as risk rises. Yet 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 at the month $\times$ front $\times$ ability quintile level (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). The averages are then standardized at the year $\times$ front level, so that (1,0) is the (x,y) coordinate of the re-centered mean, and both variables have unit standard deviations. 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 to receive the medal will be the best. As [Figure A11](#) shows, the stock of medals grew continuously throughout the war, and at any moment the medal's

signal value depends on the ability of the pilots who hold it. Initial awards may well have raised the medal’s salience and prestige, but this effect—as in the consumer good model of Pesendorfer—eventually flips sign: as progressively “worse” pilots receive the medal, its signal value falls, leaving the early, elite recipients dissatisfied with their devalued prize and potentially wanting something new.¹⁶

To examine this empirically, we rank the medalists by our Mas-Moretti measure of ability, where rank #1 is the best pilot to receive the medal. As in the Pesendorfer model, this parameter is unobservable to peer pilots and organization alike; pilots seek the medal precisely to signal their type. For each focal recipient, we study the rank of subsequent winners relative to his own: if relative rank is decreasing in the order of receipt, 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 is that many medals (77%) were ceremonially awarded on non-unique days, so award dates cannot establish a precise order. To use all recipients, we break ties with a permutation approach, randomizing the order of medals *within* non-unique dates while retaining the order *across* dates. We call this the set of permuted orderings; as the number of permutations grows, this set becomes increasingly likely to contain the true order.¹⁷

$$Y_{i,k,p} = \gamma + \beta O_{i,k,p} + \varepsilon_{i,k,p} \quad (4)$$

We fit (4) and show linear and non-parametric fits across 1,000 permutations in [Figure 7](#). For the i^{th} recipient (so $i = 1$ indexes the first medalist), we create a dataset of the relative rank $Y_{i,k,p} = \text{rank}_i - \text{rank}_k$ of the neighboring pilots (indexed by k) who received the medal in the window $[i - W, i + W]$, where the half-width W is set to 25. Stacking across medalists i and permutations p , the coefficient on relative medal order $O_{i,k,p}$ is $\beta = -0.2351$. Using a wild bootstrap that holds the award order fixed ([Appendix F](#)), we reject the one-sided null hypothesis of $\beta \geq 0$ with $p = 0.0020$. A medalist awarded medal 1 ten places after you is approximately 2.4 ranks lower.

¹⁶We assume that the exact date of award, which would reveal quality in this model, is not visible to peers. That is, having the medal is observable, but the timing of receipt is not.

¹⁷Whether 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 $\beta = -0.2725$.

The permutation approach also suggests a non-parametric means of inference: we compare the permuted orderings to a pure placebo set created by permuting medalists across *all* dates, which generates an entirely random ordering. [Figure A12](#) shows the kernel densities of the estimated slopes. The permuted slopes cluster tightly around approximately the mean recovered from the stacked estimate of (4); the placebo slopes center loosely on zero. Our empirical $p = 0.0050$ is the share of placebo slopes more negative than the mean permuted slope. The same separation, centered on a negative slope, emerges across window half-widths $W \in \{5, 10, \dots, 50\}$ ([Figure A13](#)).

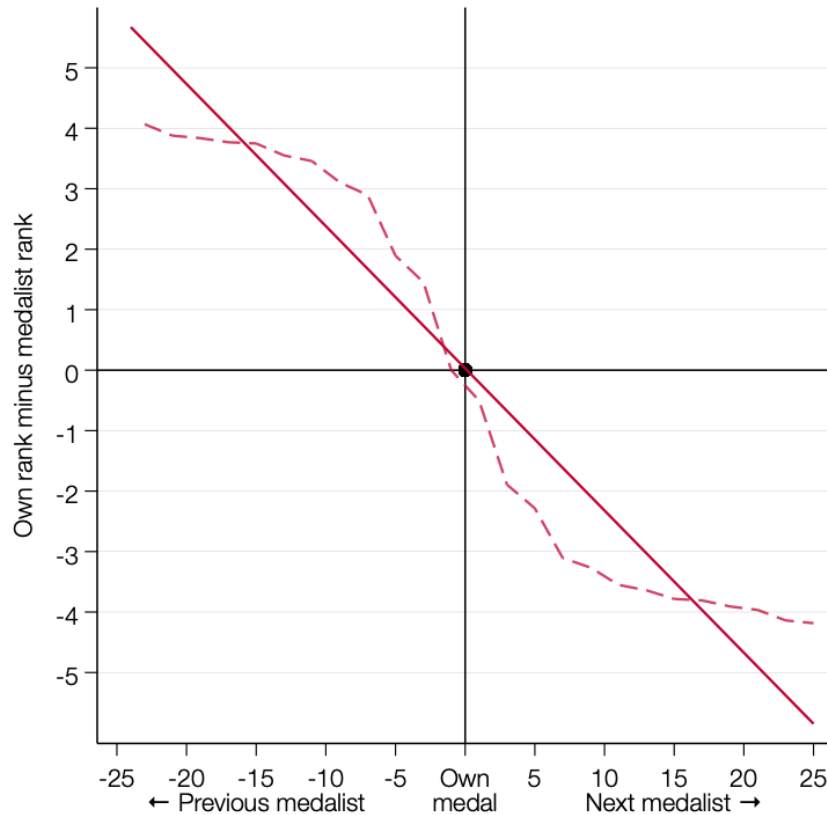
One concern with this exercise is that many victories are attained while under the influence of the medal incentives. The estimates above might then reflect a circularity: pilots who are more receptive to the medal incentives are measured as higher ability, when their out-performance is instead explained by effort. We assess this concern with a simple exercise. First, we re-estimate our Mas-Moretti ability measure excluding all pilot-months when the pilot is in-zone of a medal quota. Second, we re-estimate the cheapening coefficient β as above. Third, we compare the new estimate to the expected β when the same number of months is dropped at random. The third step is essential as, by excluding a large number of victories from the estimation sample, β suffers from classical attenuation bias that is unrelated to the circularity concern. Random drops reproduce 66% of the decline in β under the exclusion. Correcting for this attenuation, the cheapening coefficient under the purged ranking is $\beta = -0.1423$, retaining 61% of the magnitude of the headline estimate, and the implied circularity component is not statistically significant (one-sided $p = 0.1389$). We present a fuller description and results of this exercise in [Appendix F](#).

In a Pesendorfer-style model of trickle-down diffusion, high-status individuals are the first to adopt. In our setting, we must rule out a mechanical reason for the “cheapening” of medals—the well of talent running dry, which would make later awardees less outstanding by construction. [Figure A14](#) plots the ability of new pilots (entrants in their first month) at the 10th, 50th, and 90th percentiles in each month of the war: there are no time trends in entrant ability. The cheapening we document is therefore an active design choice to maintain the medal’s signaling value, not a passive consequence of the talent pool.

From an individual pilot’s perspective: those who received the medal before you were, on average, more highly skilled than you, making the medal a desirable target—you joined a highly exclusive club that will only just let you in. Those who followed, while still elite, are

lower-ranked, diluting the prestige of your medal as time goes on. In this sense, medals share characteristics with fashion trends, whose adoption initially raises demand before interest declines as “too many” lower-ranked types adopt the style—leading pilots and fashionistas alike to chase the next status symbol.

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



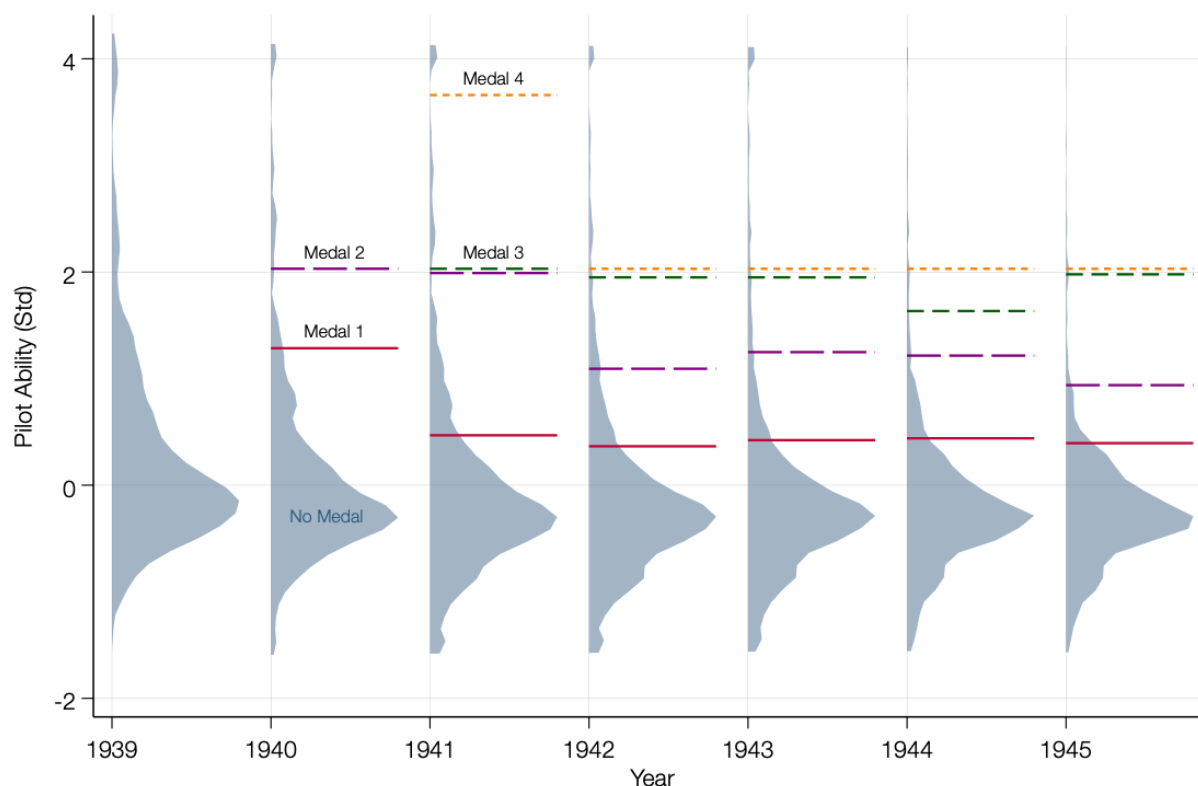
Notes: Figure shows the linear and local polynomial fit of relative rank and relative medal order for the 412 pilots with a dated medal 1 award (418 pilots ever held it), permuted 1,000 times to resolve the ambiguous order of same-day medalists. Rank is in terms of ability among all pilots who are awarded 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 the x^{th} medalist since the focal pilot was awarded the medal.

As a result of progressive cheapening, recipient quality declines over time. [Figure 8](#) plots, for each year of the war, the distribution of Mas-Moretti ability scores for pilots without any Knight’s Cross medal, marking the median ability of medal holders at each level. Between 1940 and 1941, the average quality of medal 1 holders falls sharply; between 1941 and 1942, medal 2 holders show a similar decline. When medal 4 was introduced in 1941, its holders’

median quality was markedly higher than medal 3's; just one year later, it had fallen to barely above medal 3.

Overall, the German air force managed an elaborate system in which quota changes and medal introductions counteracted the natural downward “quality drift” of the Pesendorfer model. The separation of pilots of different abilities was broadly achieved, keeping medals useful signals despite the tendency for recipient quality to decline.

Figure 8: Quality of medal recipients declines over time

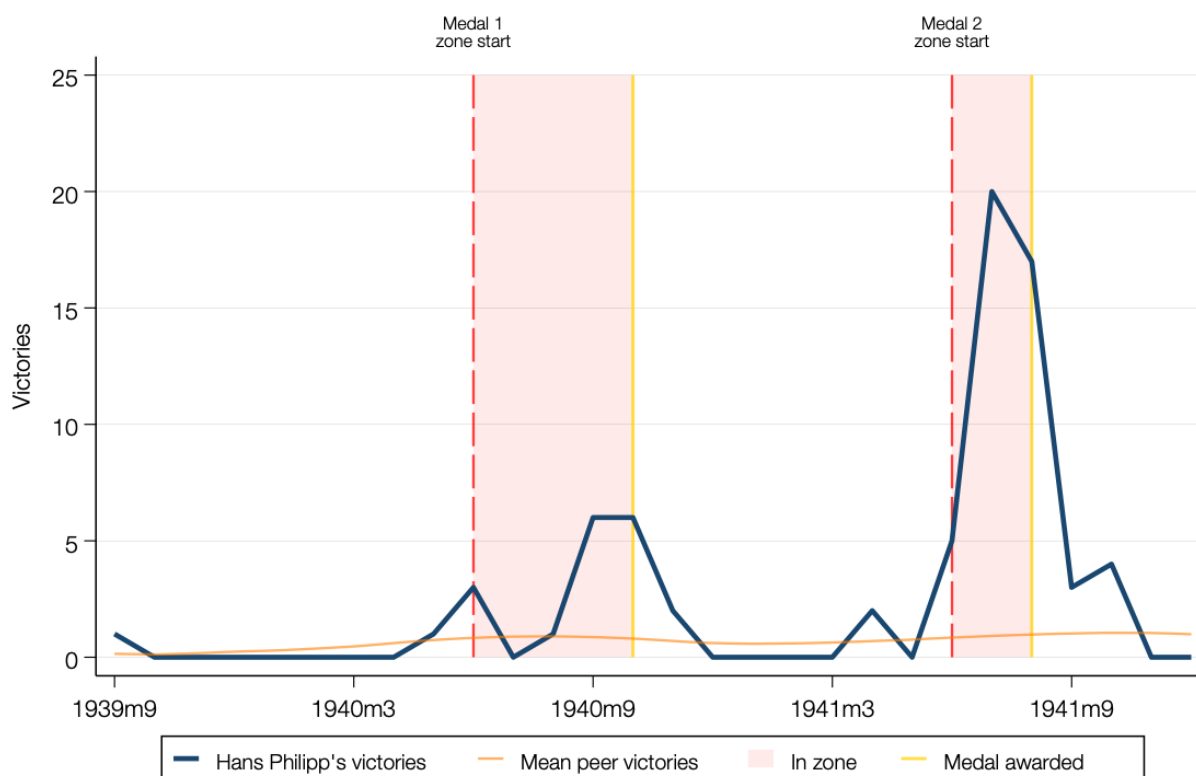


Notes: Figure shows the distribution of ability scores (Mas and Moretti, 2009) for each year for pilots who do not hold any medal. Horizontal lines indicate the median quality of medal recipients at each level of the Knight's Cross, from medal 1 through medal 4, once available.

4 Incentive Effects of the Medal “Ladder”

Did the prospect of higher medals generate a similar increase in effort as the initial award? Consider Hans Philipp (Figure 9), one of the top aces in our data, credited with over 200 career victories. The dark blue line shows his monthly victory tally; the orange line, the average for his squadron peers; dashed red lines mark where the “zones” for medals 1 and 2 begin; gold lines, the awards. Philipp’s performance accelerates sharply—both in absolute terms and relative to his teammates—as he enters the qualifying range for each medal, and slumps back toward his peers’ level immediately after each award.

Figure 9: Quotas and performance: case study evidence



Notes: Figure shows the performance of fighter ace Hans Philipp (b. 1917) over time. The dark blue line gives the monthly victory score. The orange line gives the mean score of squadron peers. Dashed red vertical lines mark the pilot’s entry into the zone of a medal quota (fifteen victories below it); shaded periods are in zone, from entry until the medal is awarded. The gold lines indicate when a medal was awarded.

To go beyond anecdotal evidence, we estimate the stacked event study in Figure 10, using

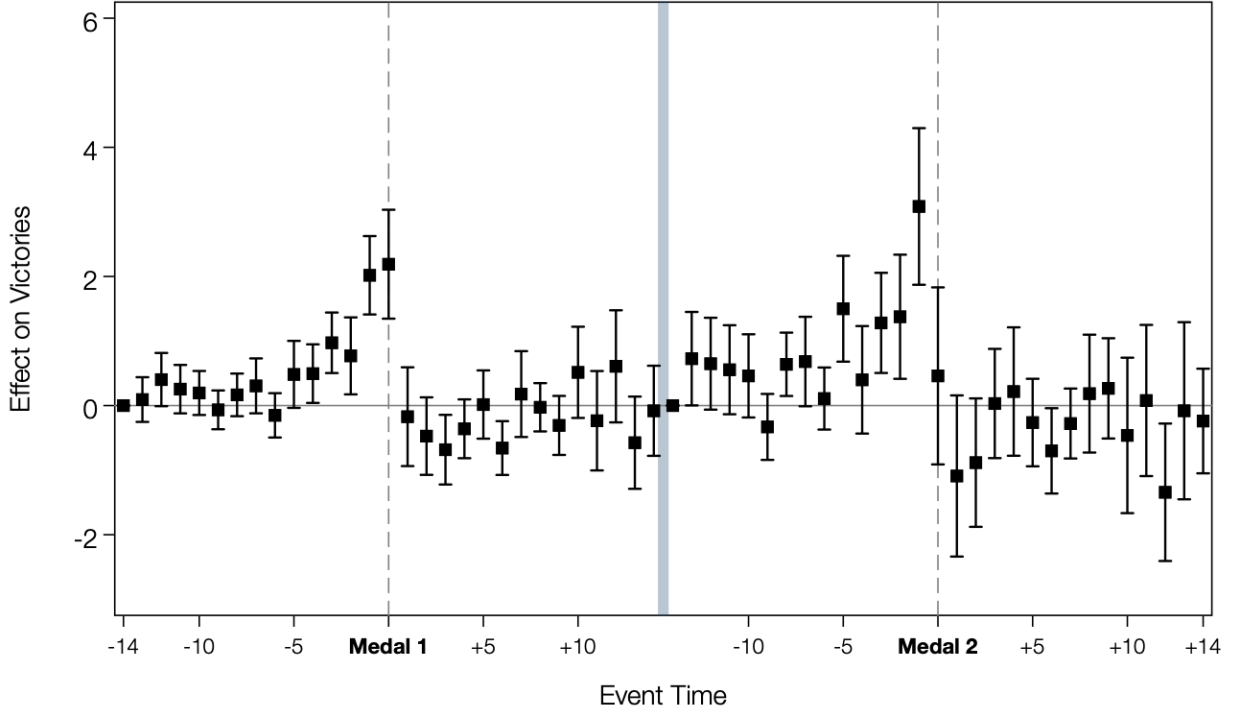
(5)—a version of (2) for two events, meeting the quotas for medals 1 and 2, indexed $m = 1, 2$.¹⁸ Each event is estimated on its complete window of 14 leads and lags, with periods outside the window absorbed into binned pre- and post-window indicators and $j = -14$ as the leave-out period, and the two event windows are stacked: the panel enters once per event, with event-specific pilot and period fixed effects and covariates $\mathbf{X}_{i,t}$ (front indicators and one lag of victories). A pilot-month within 14 months of both quota crossings therefore enters both event windows; we cluster standard errors at the squadron level. We adjust for the mechanical bias as in Section 2, with 25 placebo quotas per pilot and medal (omitting the placebo stacking from the notation).

$$Y_{i,t,m} = \alpha_{i,m} + \delta_{t,m} + \sum_{j \in \mathcal{J}} \mu_{j,m} \mathbb{1}\{t - E_{i,m} = j\} + \mathbf{X}'_{i,t} \boldsymbol{\beta}_m + \varepsilon_{i,t,m}, \quad m \in \{1, 2\} \quad (5)$$

Pilot effort rises again in the run-up to the second medal quota and falls back once it is met. Summed over the six months to the quota, the increase is 6.93 (s.e. 0.61) victories for medal 1 and 8.10 (s.e. 1.14) for medal 2; the difference of 1.17 is not statistically distinguishable from zero ($p = 0.31$). The timing differs slightly: the medal 2 run-up begins earlier and peaks in the month before the quota is met, while for medal 1 the increase is strongest in the final two months. This collapse-and-reset profile is difficult to reconcile with a standard elimination tournament, in which the option value of remaining prizes keeps incentives monotonic (Rosen, 1986). Table A3 estimates the static version of the design and extends the analysis to the third medal: though the limited number of awards leaves us underpowered, the third in zone treatment exhibits the same pattern in the OLS estimates. The table also reports pseudo-Poisson maximum likelihood estimates, given the count outcome (Cameron and Trivedi, 2015), which confirm the pattern for the first two medals.

¹⁸We plot the raw data for medals 2 and 3 (in the fashion of Figure 2) in Figure A15.

Figure 10: Pilot performance spikes repeatedly at each medal quota



Notes: Estimated effects on victories per month as a function of time relative to meeting the quota for medal 1 (left, 284 quota crossings) and medal 2 (right, 161). The event study model (5) estimates each event on its complete window of 14 leads and lags, with the 14th lead as the reference period and periods outside the window absorbed into pre- and post-window indicators; the two event windows are stacked with event-specific pilot and period fixed effects, so a month within both windows enters both panels. Estimation via OLS with controls for front and one lag of victories. Adjustments are made to remove the mechanical bias discussed in Section 2.1, with 25 placebo quotas per pilot and medal. 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” increase their effort—a pattern that repeats for each new medal on the ladder. Reaching the zone, however, requires effort and ability beyond those of most pilots. Who takes part in this rat race, and when do they begin? Medal ladder extensions provide discrete moments that might incentivize any medal-minded pilot to exert effort—even one far from qualifying for the new medal. This would be in line with a salience model, in which new awards highlight the status benefits of medals relative to the risk (Bordalo et al., 2012). Alternatively, the introductions might be moments of demotivation—the top keeps getting further and further away.

Analyzing the effect of medal introductions on pilots not in immediate “danger” of receiving the medal is challenging. New medals were introduced only a few times, and for the vast majority of pilots they were irrelevant—with little to no chance of ever meeting the quota, such pilots are in effect untreated. For the most highly skilled, by contrast, the prospect of (eventually) receiving the medal was an important source of motivation.

Ex post, we know who those pilots were: the ones who reacted markedly as they entered the “zone”. Since we are interested in the *timing* of performance changes, we restrict the analysis to this potential treatment group and compare how introductions affected performance by pilots’ prior “medal chasing”. Anecdotally, some soldiers were keener on medals than others, jokingly said to suffer from “throat ache”—Knight’s Crosses were worn around the neck (Hartmann, 2009). 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, 2025; Liu et al., 2024).

The first step is a measure of medal sensitivity. We use the causal forest method, which builds on the causal trees of Athey and Imbens (2016), as implemented by Yadlowsky et al. (2025), to calculate conditional average treatment effects (CATEs) of being in zone of the medal 1 quota, with noble birth and officer status as heterogeneity-inducing covariates. The targeting operator characteristic in Figure A16 shows substantial heterogeneity: the most medal-sensitive pilots have treatment effects 0.16 victories per month above the unconditional average. 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). We estimate an interacted fixed effects model on the control group to recover $r = 1$ ¹⁹ unobserved time-varying factors, infer treated units’ factor loadings from their pre-treatment outcomes, and use factors and loadings to predict counterfactual post-treatment outcomes. The model also includes two-way fixed effects and a combat front indicator. Intuitively, each “medal-sensitive” pilot is matched to a synthetic control of pilots with similar pre-introduction trajectories, and the average treatment effect on the treated (ATT) is the gap between actual and imputed outcomes.

Relative to standard two-way fixed effects, this approach flexibly handles unobserved heterogeneity through latent factors, under the strong assumption that the factors apply equally

¹⁹Results are similar or larger for $r = 2-4$.

to control and treatment units, and it exploits our long pre-treatment period to estimate the loadings precisely.²⁰ Within-period comparisons also sidestep the endogeneity of introduction timing to aggregate combat conditions, which apply equally to medal-minded and other pilots. The estimator requires no leave-out period, so ATTs reflect the *absolute* difference between medal-minded and control pilots; for a *relative* interpretation comparable to our other results, we re-center estimates around a reference period.

“Medal-minded” pilots are overwhelmingly *not* recipients of the higher medals whose introduction we study: fewer than one in 4 recipients of medal 1 go on to win medal 2, and only one in 14 win medal 3. We exclude, in each case, pilots who later win the medal being introduced, and pilots “topped out” at the previously top medal, to avoid measuring a cheapening effect of the new medal on the old (the focus of the previous section). The effects can then be interpreted as the pull of the new award on pilots still some distance from qualifying. With these restrictions, we study 338 medal-minded pilots.

Figure 11 shows the results for the introductions of medal 2 in June 1940 and medals 3 and 4 (simultaneously) in July 1941, in a short 12-month window centered at $t = 0$ (the events are only 13 months apart). Medal-minded pilots are, at least initially, motivated by the announcement of a new, higher goal, even one far beyond their current realistic aspirations: they earn approximately 0.2 extra victories per month for two months before the effect fades to zero, though the estimates are imprecise. This is consistent with our quota evidence—pilots increase effort around salient events rather than smoothly in all periods—though the effect is noticeably smaller than in the quota exercise. We date introductions by the promulgation of the new administrative procedure, which may lag pilots’ awareness of the impending extension. Despite this, the estimate at $t = -1$ is close to zero.

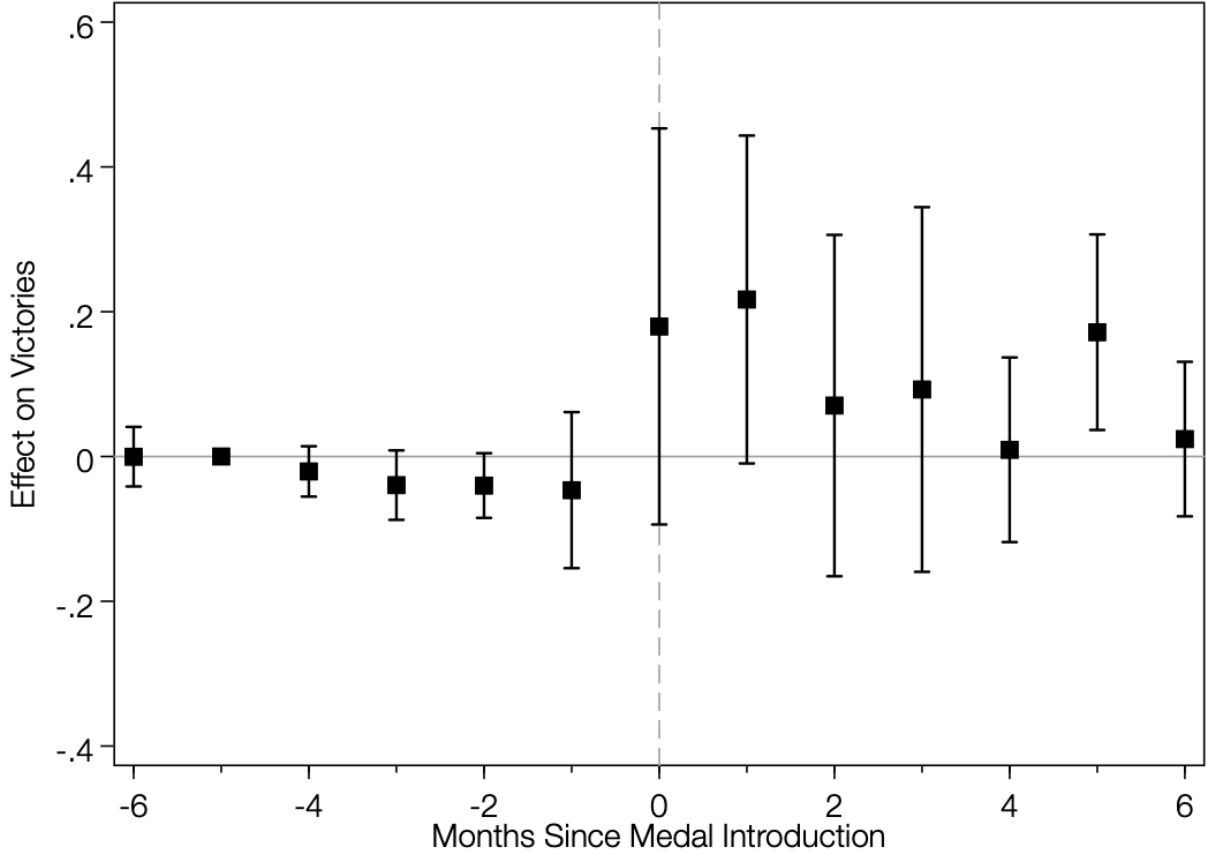
A standard robustness exercise excludes several pre-treatment periods from the model-fitting step as placebos; their estimated ATTs should be insignificant if the model is correctly specified, alleviating concerns of over-fitting to the pre-period. Figure A17 confirms that the estimator performs well.

These results support an organization’s ability to repeatedly extract effort through a hierarchical system of incentives: pilots respond to the creation of higher-order incentives even when the targets are remote, suggesting awareness of the entire medal system, not just the

²⁰Treated pilots with fewer than five pre-treatment observations are dropped from this analysis.

rungs immediately relevant to their careers. A limitation is the lack of true cross-sectional variation in exposure, which necessitates within-period comparisons against pilots we assume 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, $t = -5$. 95% confidence intervals from 500 bootstrap draws are indicated by the capped spikes.

5 Mechanisms and Magnitudes

This section presents additional, indirect evidence for our proposed mechanism—status competition—and derives a lower-bound estimate of the total number of additional Allied

planes destroyed because of the German medal system.

5.1 Heterogeneity of the in zone response by pilot prestige

To understand what motivates medal chasing, we examine alternative sources of prestige. Many (but far from all) medal-eligible pilots are officers with the honor of command; some are noblemen; many hold other, lower-tier medals. We hypothesize that a highly prestigious medal is more motivating for pilots with few or no other sources of social standing, in line with status signaling models (e.g., [Spence, 1973](#))—pilots with strong existing signals have weaker incentives to invest in costly additional signals.²¹

We measure a pilot’s prestige with an equally weighted index of three markers: holding (i) the German Cross (*Deutsches Kreuz*), (ii) a squadron commander’s rank of first lieutenant (*Oberleutnant*) or higher, and (iii) a noble title. Each indicator is standardized and the three are averaged, with the average rescaled to unit standard deviation.²² Examples of high-status pilots by this measure include Diethelm von Eichel-Streiber, a landed politician’s son with a decorated Spanish Civil War record, and Hubertus von Bonin, son-in-law of a Prussian general and an experienced squadron commander.

We study the interaction between being “in zone” of the medal quota (as defined above) and alternative sources of prestige, using a static version of the bias-adjusted estimator from Section 2.1: we fully interact the variables of interest with the “real” in zone indicator on the stacked data. Column 1 of [Table 2](#) replicates the static equivalent of the dynamic result in [Figure 5](#)—pilots work harder when close to the quota. Column 2 shows that alternative prestige is associated with slightly higher performance; the interaction in Column 3, however, is strongly negative: high-status pilots are less motivated by the prospect of the next medal than pilots without prestige markers. Medals are thus *substitutes* rather than *complements* for other status signals. The index result is consistent with each of its three components:

²¹In a standard signaling framework, the return to acquiring an additional signal is the increment to the audience’s posterior belief about the sender’s type, while the cost of acquisition is positive. When existing signals have already pushed the posterior toward the upper end of the type space, a further signal shifts beliefs—and thus status—by less, while its cost is unchanged. The marginal net return to a new signal is therefore decreasing in the stock of signals already held, so pilots with strong existing status markers should respond less to the prospect of a medal.

²²Noble status is time-invariant and indicated by the presence of a 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.

entering the markers one at a time and jointly (Table A4), every interaction is negative. Studying the medals separately, the result is precisely estimated for medal 1 and has the same sign, but is imprecise, for medals 2 and 3 (Table A5). These results echo Charles et al. (2009): acquiring visible status signals matters more for individuals from lower-status groups.

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

	Victories		
	(1)	(2)	(3)
Inzone=1	0.440*** (0.048)	0.429*** (0.047)	0.528*** (0.054)
Prestige		0.056* (0.032)	0.044 (0.033)
Inzone=1 × Prestige			-0.123*** (0.028)
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’s cumulative victories are no more than 15 below the quota for his next medal (medals 1–3), or at or above it, and that he has not yet been awarded that medal. Prestige is the equally weighted index of indicators for holding a squadron commander’s rank, the DK medal, and being of noble birth (each standardized, averaged, and rescaled to unit standard deviation). 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.

5.2 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 31% of their time “in zone”, but scored 46% of their victories there (Figure A18).

To arrive at an aggregate figure, we sum the area under the curve of victories as pilots approach the quota. In the raw data (Figure 2 and Figure A15), where event time is measured relative to medal receipt, this yields 11.84 victories for medal 1 and 17.67 for medal 2.²³ After correcting for mechanical bias with our estimator from Section 2, where event time is measured relative to meeting the quota, the area is 6.34 (s.e. = 0.63) per in zone medal 1 pilot.²⁴ Comparing the two areas—total victories summed across event time—54% of the raw victories bump can be causally attributed to the first medal quota; since the two areas differ in event timing and summation window, we read this share as approximate. For medal 2, the adjusted area is 8.37 (s.e. = 0.93), representing 47% of the raw total.

We then scale these causal aggregates by the number of treated pilots. Table 3 offers three approaches that yield comparable results, differing in who counts as treated. Approach 1 treats all pilots who ever hit the zone—an overestimate if risk-taking pilots exit at higher rates. Approach 2 treats only pilots who receive the medal they are in zone for—likely an underestimate, for the same reason. Approach 3 weights each event-study coefficient $\hat{\theta}_j$ in (3) by the number of pilots observed at that lag j .²⁵ In total, the quota system caused 3,100–5,800 extra victories, representing 6–11% of total victories in our dataset. Against the median pilot’s three career victories, the immediate effect of the medal system—sharper incentives as the medal loomed large—added the same number of aerial victories as deploying around 1,000–1,900 additional (median) pilots.

These figures plausibly underestimate the true total effect: by focusing on the aspects of the medal system we can causally identify, we abstract from effects on behavior outside the zone, which Section 4.1 suggests may also be significant.

²³For the raw data, the area is the sum of mean monthly victories over the award month and the six months before it ($j = -6, \dots, 0$), net of seven times the mean over the baseline window $j = -12, \dots, -8$.

²⁴For the adjusted estimates, we sum the coefficients $\hat{\theta}_j$ for $j \leq 0$ where the 95% confidence interval excludes zero.

²⁵In 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

Description	Medal 1		Medal 2		Total Inzone Victories	
	AUC	Inzoners	AUC	Inzoners	Count	Share of Grand Total
Approach 1	6.34	623	8.37	222	5,808	10.59% (0.81%)
Approach 2	6.34	418	8.37	104	3,521	6.42% (0.51%)
Approach 3	1793.83		1331.49		3,125	5.70% (0.42%)

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). In Approach 1, we scale by the count of pilots who *ever* enter the zone for a medal, regardless of completion. In Approach 2, we only scale by the number of pilots who win each medal. In Approach 3, we weight each event-study coefficient $\hat{\theta}_j$ from (3) by the number of pilots observed in the lag j . Standard errors for the final shares are in parentheses and 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, theory predicts that such a system can be created if status concerns are sufficiently strong and ability differs substantially (Moldovanu et al., 2007). We examine this question empirically in a dynamic setting, providing direct evidence from WWII that the German air force successfully created a ladder of highly coveted medals whose rungs multiplied as more workers earned the top awards—and that this system increased performance.

Using a modified event-study design and data on the monthly performance of over 5,000 German pilots, we show that performance increases markedly as pilots approach a medal’s eligibility threshold and slumps immediately after the award, suggesting a sharp change in incentives and effort. The effect is weaker for pilots who already possess other markers of high social status, in line with a standard signaling game.

As pilots approach the next medal on the ladder, the same effort path repeats itself: achieving the already highly prestigious first medal does not satisfy them. A synthetic control design shows 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 creates artificial scarcity rather than selling novelty, the underlying mechanism is related: the best pilots win the medal first, low types follow, and the resulting cheapening provides a rationale for new, rarer variants that draw high types back into the rat race.

While extensions of the medal ladder could theoretically be demotivating, our results show that a hierarchical system of status-based incentives can repeatedly elicit effort, generating a rat race among workers. The net effect was substantial—we estimate the ladder added between 3,100 and 5,800 victories to German pilots’ tally during the war (equivalent to 6–11% of the total).

Several features of our setting bound the generalization of these results. The organization’s credibility—the national armed forces during wartime—gave it leeway to construct status symbols that are *ex ante* desirable. It offered these rewards to men mostly in their 20s, an age when status-seeking is common and risk tolerance high. Stakes were mortal, and pilots had essentially no outside options—a point that cuts in an instructive direction: where workers are mobile, public recognition also reveals quality to competing employers, so organizations that fear poaching may ration recognition of their best performers. One common objection, however, does not apply. Status incentives here operated alongside tangible, non-financial career instruments—promotions show no differential timing around medal quotas ([Figure 4](#)), and medals brought some perks, such as home-front propaganda tours—so the effects we document are not an artifact of symbolic rewards being the only available margin. Not every organization will enjoy the same capabilities or so well-suited a population; the resulting variation in rat race intensity should be explored in other settings.

References

- Abadie, Alberto and Jaume Vives-i-Bastida**, “Synthetic Controls in Action,” in “Advances in Economics and Econometrics: Twelfth World Congress,” Cambridge: Cambridge University Press, 2025, pp. 195–224.
- 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.
- 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*, June 2017, *63* (6), 1657–1672.
- Ashenfelter, Orley**, “Estimating the Effect of Training Programs on Earnings,” *The Review of Economics and Statistics*, 1978, *60* (1), 47–57.
- Ashraf, Nava and Oriana Bandiera**, “Social Incentives in Organizations,” *Annual Review of Economics*, 2018, *10* (1), 439–463.
- , —, 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.
- Auriol, Emmanuelle and Régis Renault**, “Status and incentives,” *The RAND Journal of Economics*, 2008, *39* (1), 305–326.
- 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, Philip and Donald T. Campbell**, “Hedonic Relativism and Planning the Good Society,” in Mortimer H. Appley, ed., *Adaptation-Level Theory: A Symposium*, New York: Academic Press, 1971, pp. 287–302.
- 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.
- 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.
- 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.

- 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.
- Gallus, Jana**, “Fostering Public Good Contributions with Symbolic Awards: A Large-Scale Natural Field Experiment at Wikipedia,” *Management Science*, 2017, *63* (12), 3999–4015.
- 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.
- 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 of *Quellen und Darstellungen zur Zeitgeschichte*, Munich: Oldenbourg, 2009.
- Heath, Chip, Richard P. Larrick, and George Wu**, “Goals as Reference Points,” *Cognitive Psychology*, February 1999, *38* (1), 79–109.

- Hopkins, Ed and Tatiana Kornienko**, “Running to Keep in the Same Place: Consumer Choice as a Game of Status,” *American Economic Review*, 2004, *94* (4), 1085–1107.
- 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 Madarász**, “Superiority-Seeking and the Preference for Exclusion,” *The Review of Economic Studies*, July 2024, *91* (4), 2347–2386.
- 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,” *American Economic Review*, 1996, *86* (3), 329–348.
- Lazear, Edward P. and Sherwin Rosen**, “Rank-Order Tournaments as Optimum Labor Contracts,” *Journal of Political Economy*, 1981, *89* (5), 841–864.
- 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.
- 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.

- Murray, Williamson**, *Strategy for Defeat: The Luftwaffe, 1933–1945*, Maxwell Air Force Base, AL: Air University Press, 1983.
- 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 Hoffmann, 1966.
- 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,” *American Economic Review*, 1995, 85 (4), 771–792.
- Prendergast, Canice**, “The Provision of Incentives in Firms,” *Journal of Economic Literature*, March 1999, 37 (1), 7–63.
- Rayo, Luis**, “Monopolistic Signal Provision,” *The B.E. Journal of Theoretical Economics*, January 2013, 13 (1), 27–58.
- Rosen, Sherwin**, “Prizes and Incentives in Elimination Tournaments,” *American Economic Review*, 1986, 76 (4), 701–715.
- Sales & Marketing Management**, “Windows Around the World,” September 2014. <https://salesandmarketing.com/windows-around-world/>.
- Salesforce**, “Showcase Your Expertise with New Trailhead Ranger Ranks,” The 360 Blog September 2022. <https://www.salesforce.com/blog/showcase-your-expertise-with-new-trailhead-ranger-ranks/>.
- Shapley, Lloyd S.**, “A Value for n -Person Games,” in Harold W. Kuhn and Albert W. Tucker, eds., *Contributions to the Theory of Games II*, number 28. In ‘Annals of Mathematics Studies.’, Princeton: Princeton University Press, 1953, pp. 307–317.

- Simkin, Mikhail V and Vwani P Roychowdhury**, “Theory of Aces: Fame by Chance or Merit?,” *Journal of Mathematical Sociology*, 2006, *30* (1), 33–42.
- Spence, Michael**, “Job Market Signaling,” *The Quarterly Journal of Economics*, 1973, *87* (3), 355–374.
- van Creveld, Martin**, *Fighting Power: German and US Army Performance, 1939-1945*, Westport, Conn.: Greenwood Press, 1982.
- 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.
- Yadlowsky, Steve, Scott Fleming, Nigam Shah, Emma Brunskill, and Stefan Wager**, “Evaluating Treatment Prioritization Rules via Rank-Weighted Average Treatment Effects,” *Journal of the American Statistical Association*, 2025, *120* (549), 38–51.

Appendix for

Never Enough: Dynamic Status Incentives and the Rat Race

Leonardo Bursztyn

Ewan Rawcliffe

Hans-Joachim Voth

For Online Publication

List of Tables

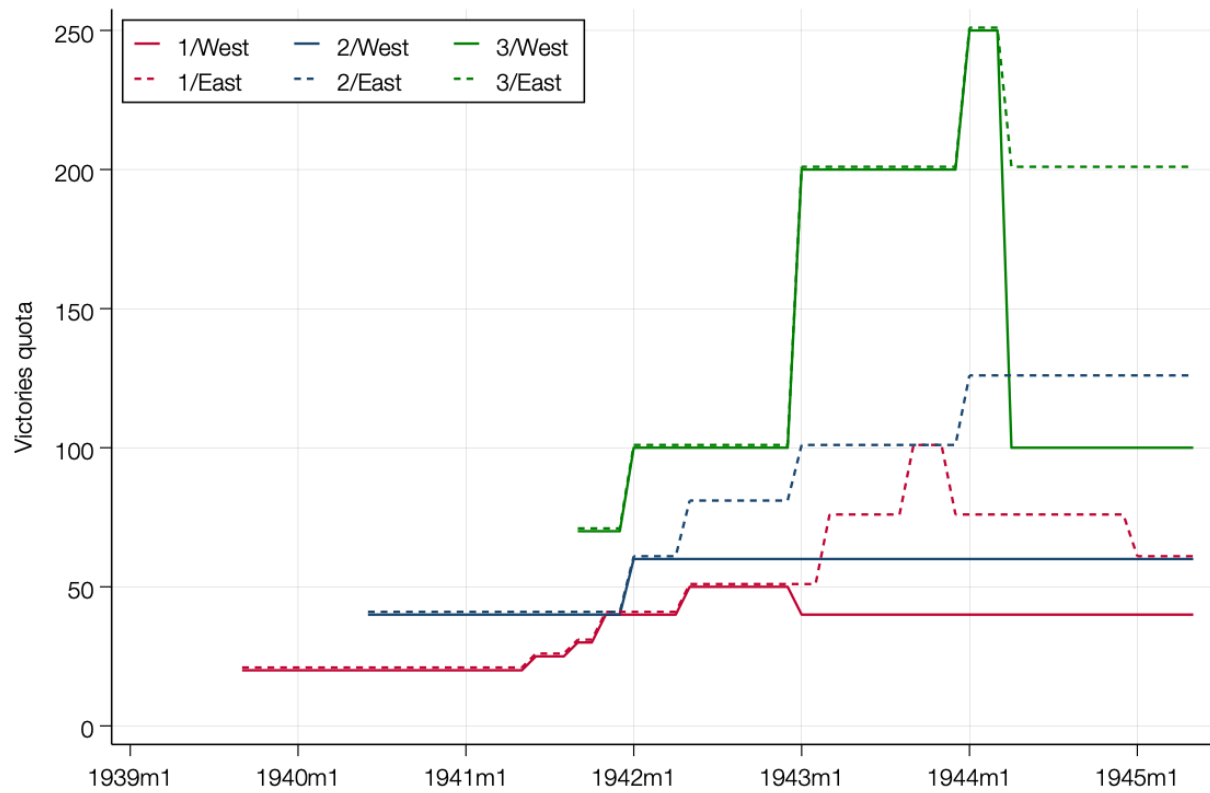
A1	Summary statistics for key analysis variables	A-5
A2	Using quota increases as an instrument for the in zone response	A-14
A3	Additional estimates of the in zone effect by medal	A-20
A4	In zone response heterogeneity: prestige index and its components	A-23
A5	Heterogeneity of the in zone response for higher medals	A-24
A6	Balance of pilot attributes based on inclusion in ability estimation	A-30
A7	Coalition value function: complementarity in squadron output	A-35
A8	Cheapening slope: attenuation check for the exclusion of in-zone months . .	A-42

List of Figures

A1	Medal quotas are adjusted throughout the war	A-3
A2	The Knight's Cross medal	A-4
A3	Medals are awarded according to an informal victories quota	A-6
A4	Process of medal award	A-7
A5	Comparison of naive and adjusted estimators in synthetic data	A-8
A6	Pilot exit risk does not track quota distance	A-9
A7	Victory rates just above and below the quota, by medal status	A-10
A8	Residualized victories track combat sorties	A-11
A9	Effort margins in the flight logs around the medal quota	A-12
A10	Pilot effort spikes around the medal quotas, robustness exercises	A-13
A11	The stock of medals awarded rises throughout the war	A-15
A12	The estimated rate of cheapening is robust to permutations in medal order .	A-16
A13	The estimated rate of cheapening is robust to the width of the comparison .	A-17
A14	Ability of entrant pilots over time is constant	A-18
A15	Raw victory rates over higher medal event time	A-19
A16	In zone treatment effects vary across pilots	A-21
A17	Pilot effort increases when new medals are introduced: placebo	A-22
A18	Time "in zone" contributes disproportionately to overall victory tally	A-25
A19	Pilot ability estimation: squadron composition variability	A-27
A20	Pilot ability estimation: peer combination frequency	A-28
A21	Pilot ability estimation: peer fixed effects and model fit	A-29
A22	Shapley decomposition of squadron output	A-36
A23	Local polynomial estimates of zone entry	A-37
A24	Quota changes and performance: case study evidence	A-39

A Additional Figures and Tables

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



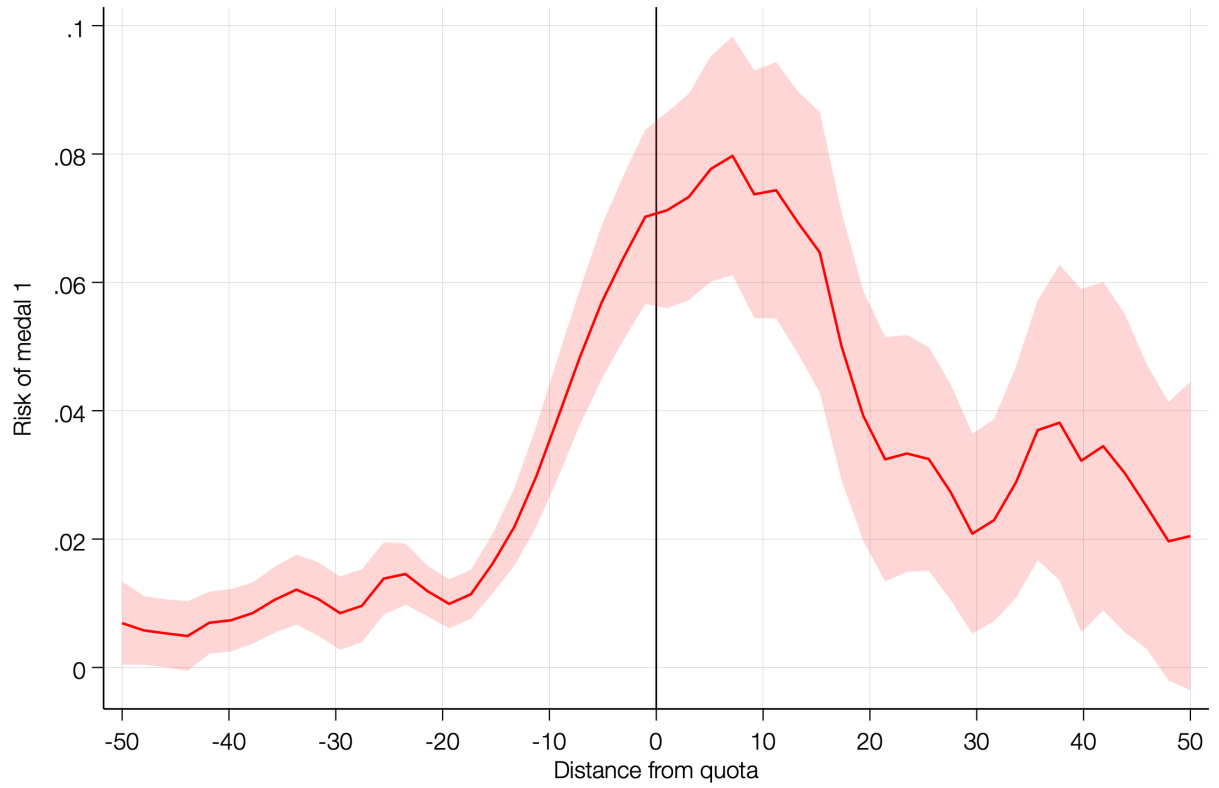
Notes: An original *Ritterkreuz des Eisernen Kreuzes* medal (medal 1). The most common variant of the medal was manufactured by Steinhauer & Lück and measured approximately 48 by 54 millimeters. We manually pixelate the swastika at the center of the medal, in line with German legal requirements.

Table A1: Summary statistics for key analysis variables

	N	mean	s.d	min	p50	p90	max
<i>A. Pilot-month level</i>							
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
Holds DK Medal (0/1)	82,348	0.11	0.31	0	0	1	1
Distance to Medal 1 Quota	82,348	-33.66	25.15	-100	-37.00	-13.00	303
<i>B. Pilot level</i>							
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	3,972	0.01	1.29	-6.20	-0.19	1.18	18.07
Total Experience (months)	5,081	16.21	17.52	1	9.00	46.00	69
Nobility (0/1)	5,081	0.01	0.12	0	0	0	1
Receives DK Medal (0/1)	5,081	0.11	0.31	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

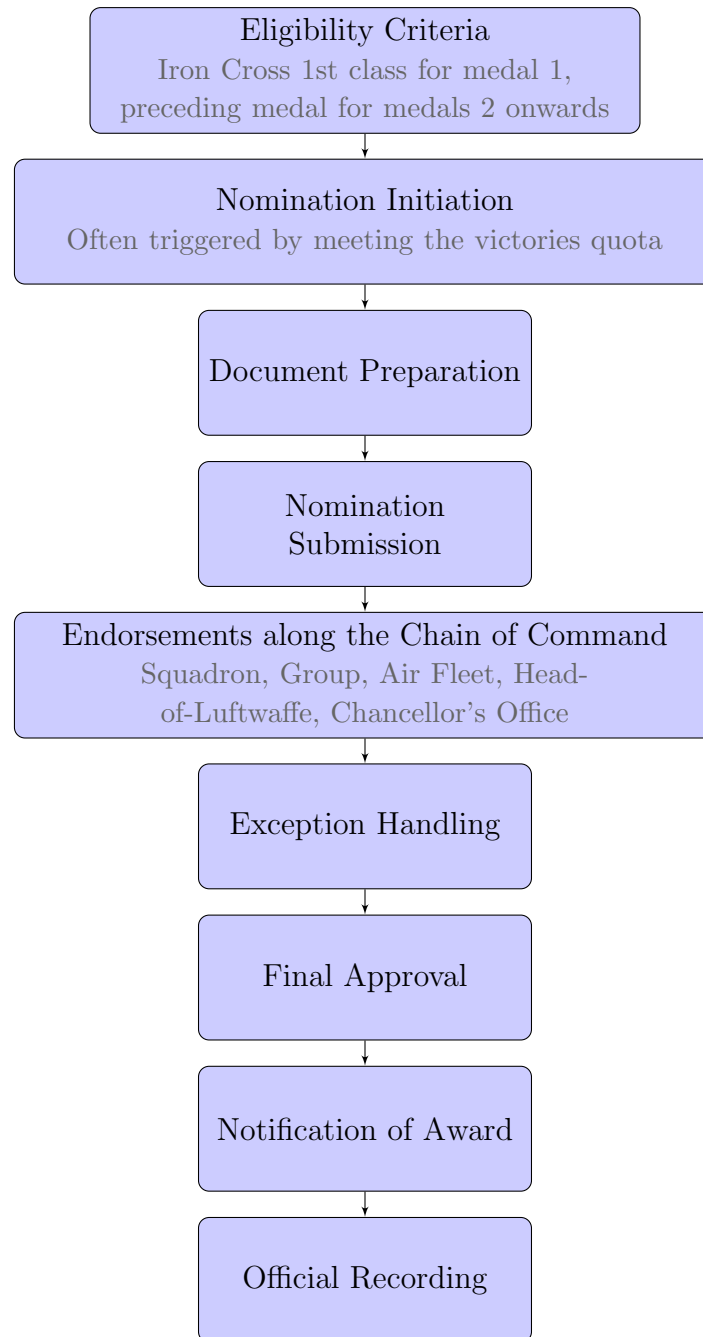
Notes: Table presents summary statistics for key analysis variables. Exit indicates either death, capture, or permanent disability. Pilot ability is the [Mas and Moretti \(2009\)](#) measure, net of peer effects. DK is the Deutsches Kreuz award. Nobility refers to all noble titles in our records (von, Freiherr, Graf, Baron, etc.). Medal 1 is the Knight's Cross. Medal 2 is the Knight's Cross with Oak Leaves; medal 3 is the Knight's Cross with Oak Leaves and Swords. Experience is the time from first appearance in the records of the Kracker Archive.

Figure A3: Medals are awarded according to an informal victories quota



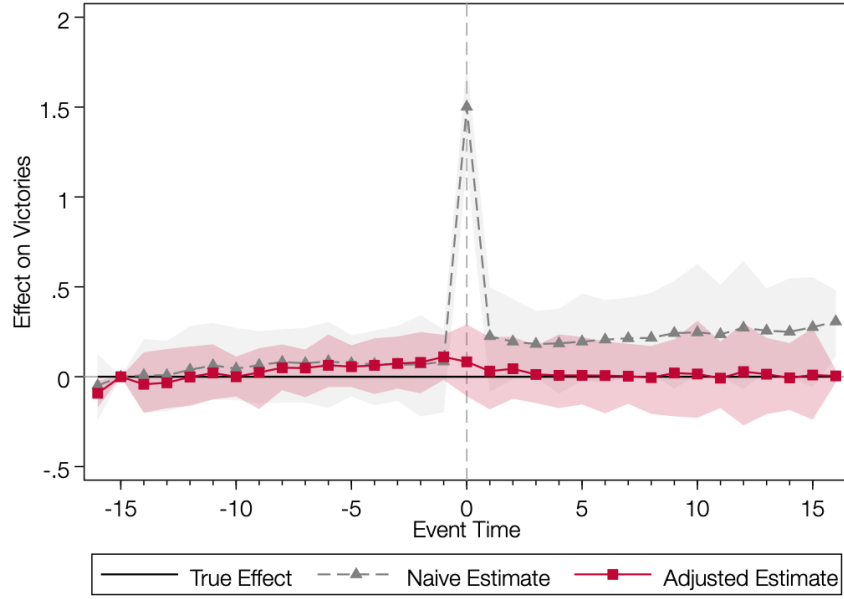
Notes: Figure shows the relationship between distance to the informal victories quota and the risk of being awarded medal 1. The horizontal axis shows the difference between a pilot's cumulative victories and the informal victories quota for the medal, which varies by period and front. The red line gives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

Figure A4: Process of medal award

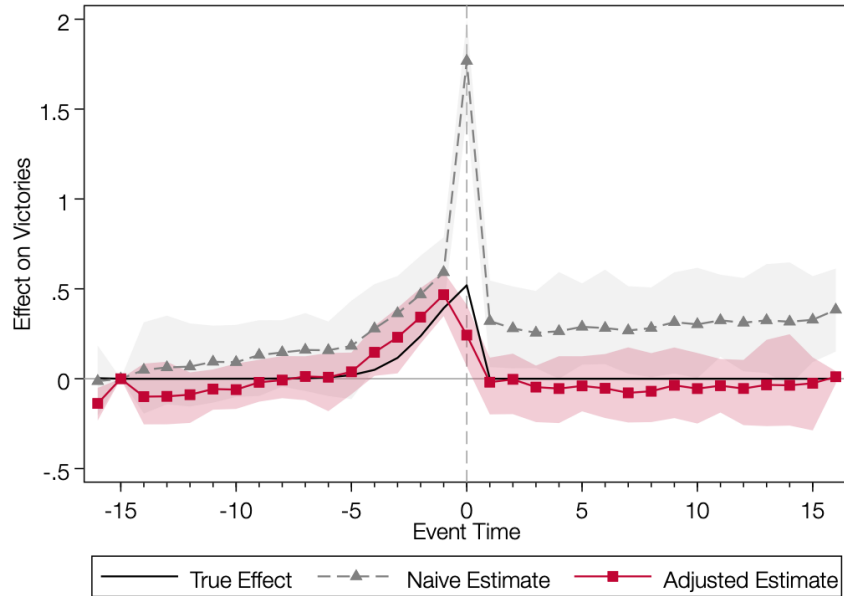


Notes: The process of qualification, nomination, and notification for the Knight's Cross medals, adapted from [Williamson and Bujeiro \(2004\)](#).

Figure A5: Comparison of naive and adjusted estimators in synthetic data



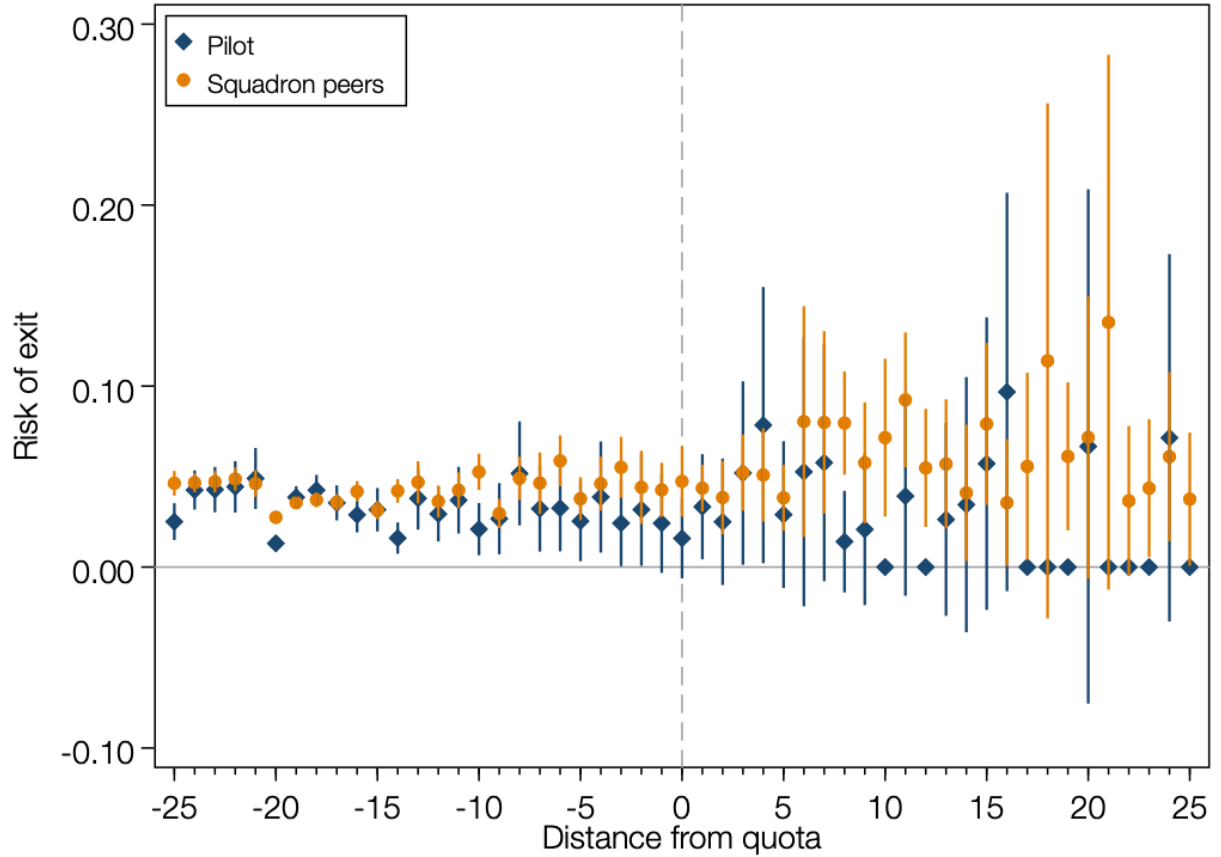
(a) Simulated effect under the null hypothesis ($\alpha = 0$)



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

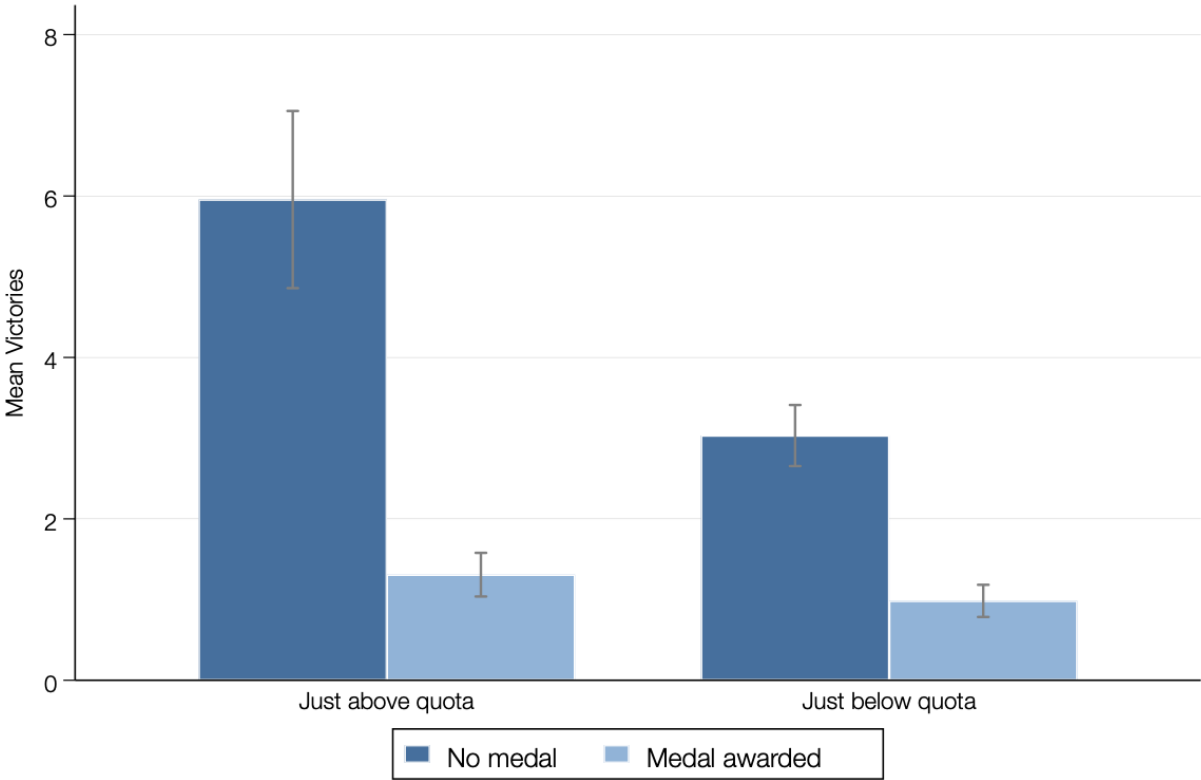
Notes: Figure shows the mean event study coefficients as estimated in 50 Monte Carlo simulations of the DGP described in [Appendix C](#), where each simulation generates $P = 25$ placebo quotas per pilot for the mechanical-correction procedure. In the upper panel, the DGP sets the true medal anticipation 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. 90% percentile confidence intervals for Monte Carlo simulation coefficients are given in the shaded areas.

Figure A6: Pilot exit risk does not track quota distance



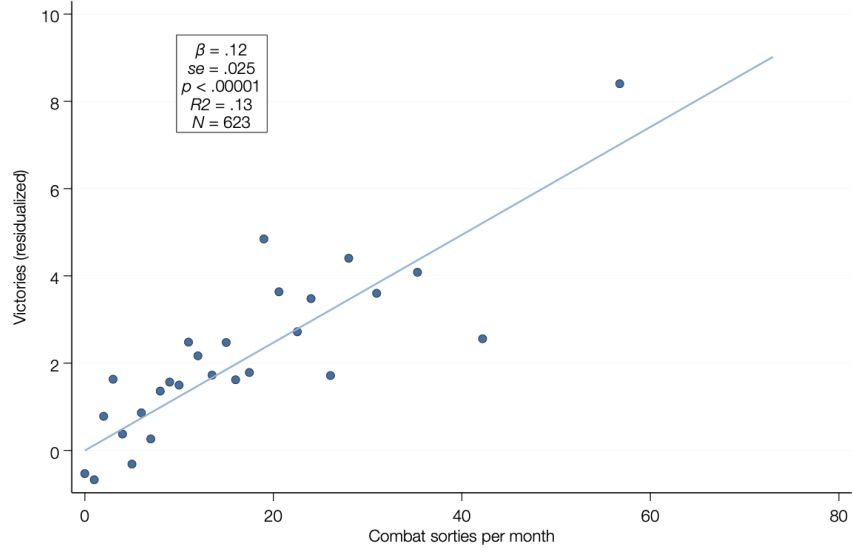
Notes: Figure shows the relationship between distance to the informal victories quota and the risk of exit. The horizontal axis shows the difference between a focal pilot's cumulative victories and the informal victories quota for the medal, which varies by period and front. The blue markers show this pilot's risk of exit conditional on distance. The orange markers show the risk of exit for peer pilots who fly in the same squadron and in the same month as the focal pilot. The means are bounded by 95% confidence intervals.

Figure A7: Victory rates just above and below the quota, by medal status

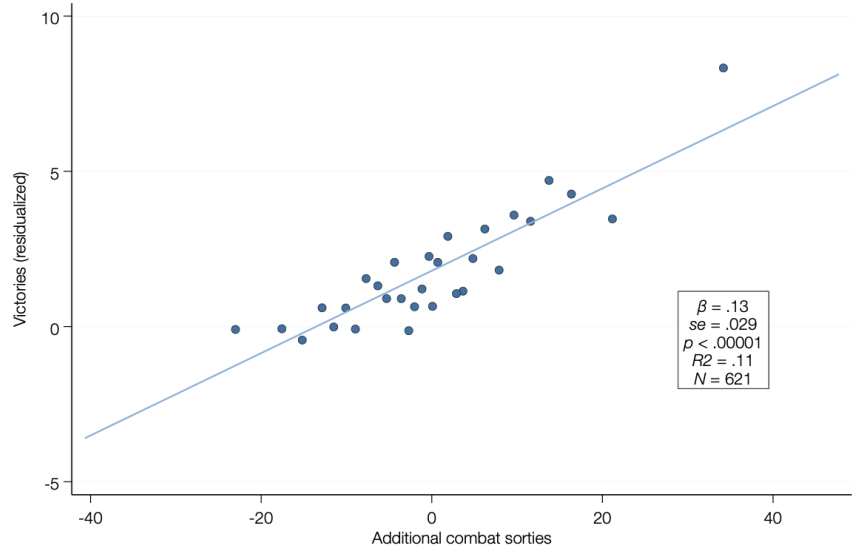


Notes: Each bar gives an average victory rate while within five victories of the medal quota, above and below. Light blue is for pilots who hold the medal corresponding to the quota, dark blue for those who do not yet have it.

Figure A8: Residualized victories track combat sorties



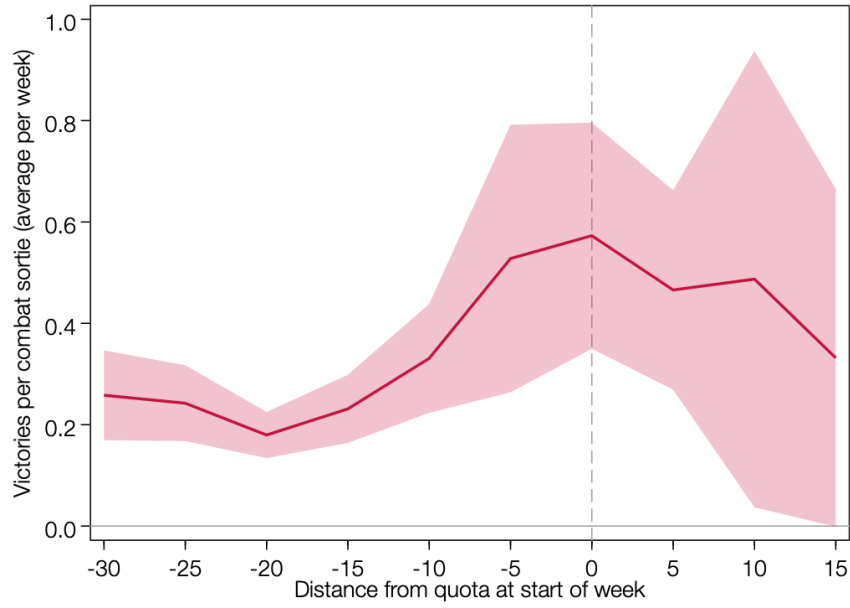
(a) Combat sorties



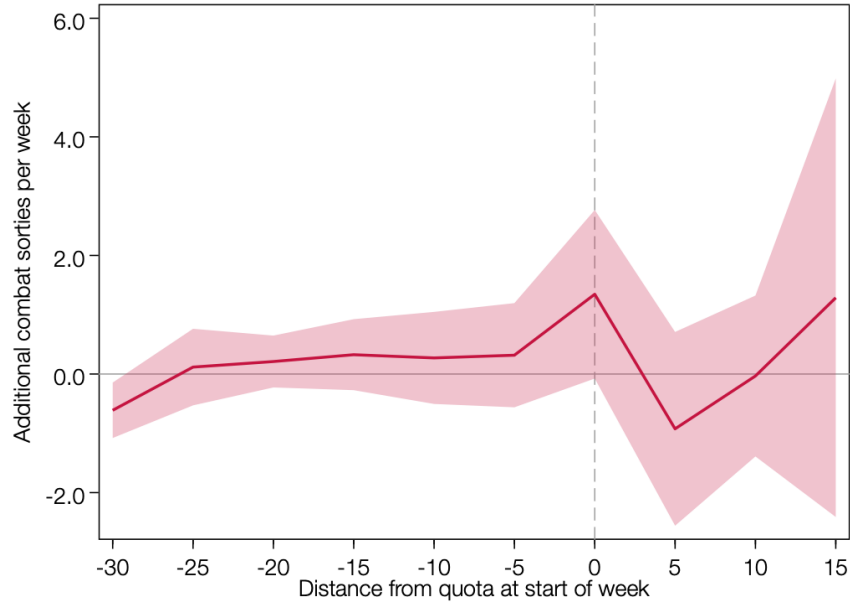
(b) Additional combat sorties

Notes: Figure shows binned scatter plots at the pilot-month level for the subsample of pilots with surviving flight logs (Ager et al., 2022). The y -axis gives monthly victories residualized on pilot, period $\times$ front, and squadron fixed effects, estimated in the full panel. The x -axis gives combat sorties flown (panel a) and combat sorties minus the leave-one-out average across all logged pilots active in the same calendar month (panel b). Blue lines give linear fits. Standard errors clustered at the level of fighter squadrons.

Figure A9: Effort margins in the flight logs around the medal quota



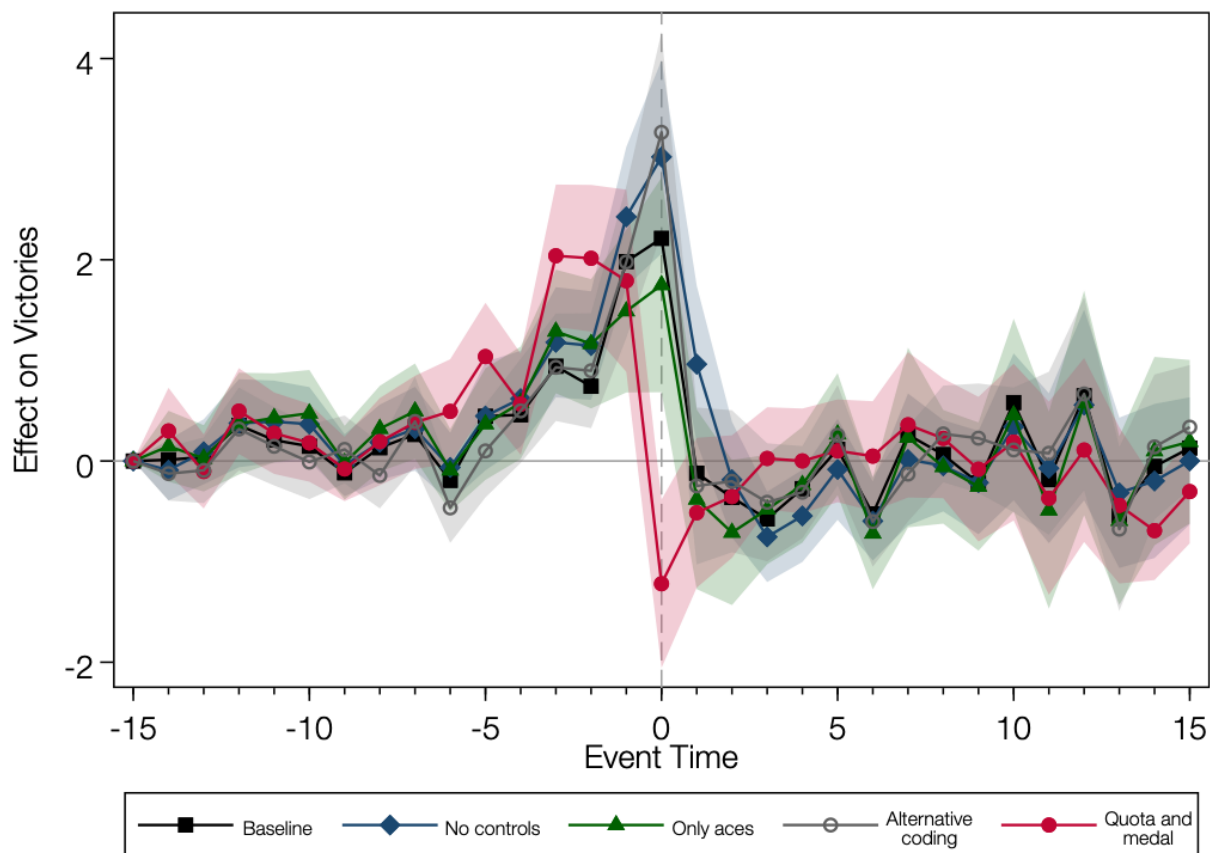
(a) Intensive margin



(b) Extensive margin

Notes: Figure shows two margins of effort at the pilot-week level for the subsample of pilots with surviving flight logs (Ager et al., 2022), as a function of the distance in victories from the next medal quota at the start of the week (binned by 5). Panel (a) shows mean victories per combat sortie. Panel (b) shows the pilot's combat sorties minus the leave-one-out average across all logged pilots active in the same week. Shaded areas indicate 95% confidence intervals. The dashed line marks the quota.

Figure A10: Pilot effort spikes around the medal quotas, robustness exercises



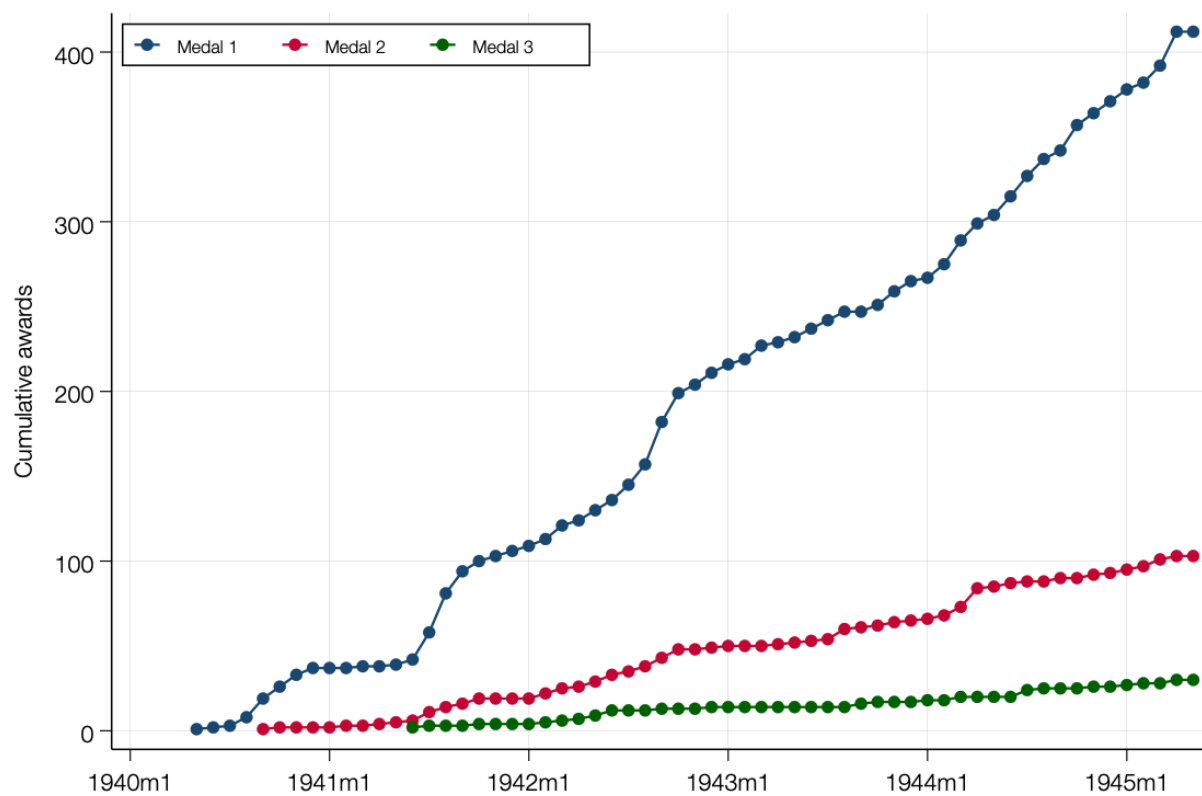
Notes: Estimated average treatment effects on victories per month as a function of their time relative to meeting the victories quota. The event study model (variants of (3)) includes 15 lags and leads of treatment, with the 15th lead ($j = -15$) as the reference period; periods further from the quota are absorbed into separate end-bin indicators. 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 pilots. 95% confidence intervals for coefficient estimates are given in the shaded areas.

Table A2: Using quota increases as an instrument for the in zone response

	Victories					
	(1)	(2)	(3)	(4)	(5)	(6)
In Zone	1.477*** (0.119)	0.990*** (0.091)		2.750*** (0.801)	3.162** (1.430)	4.544*** (1.235)
Lag Victories		0.311*** (0.012)				0.261*** (0.021)
Awarded KCR 0/1		-0.338*** (0.098)				-0.385*** (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)

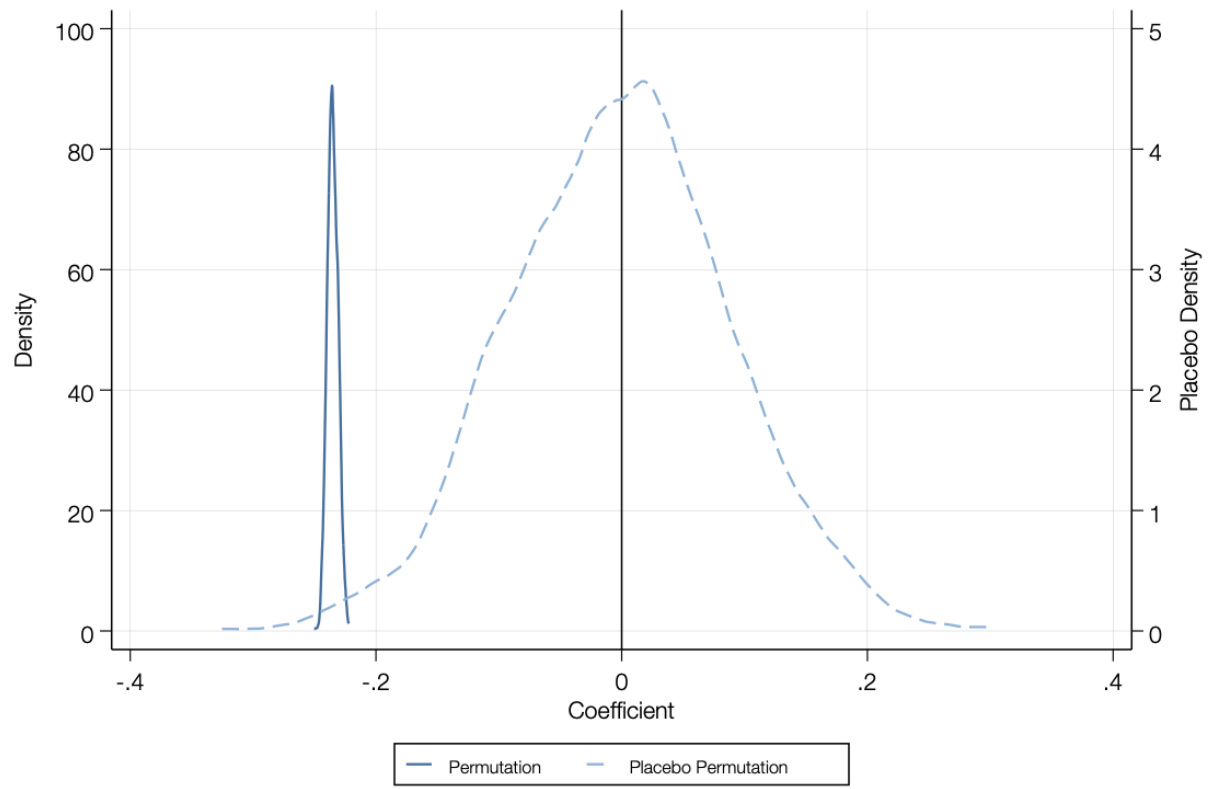
Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The table presents regression estimates. The outcome is individual monthly victories. Inzone indicates pilots whose cumulative victories are no more than 15 below the quota for their next medal, or at or above it, and who have not yet been awarded that medal. 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 $\times$ period $\times$ 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 $\times$ 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 A11: The stock of medals awarded rises throughout the war



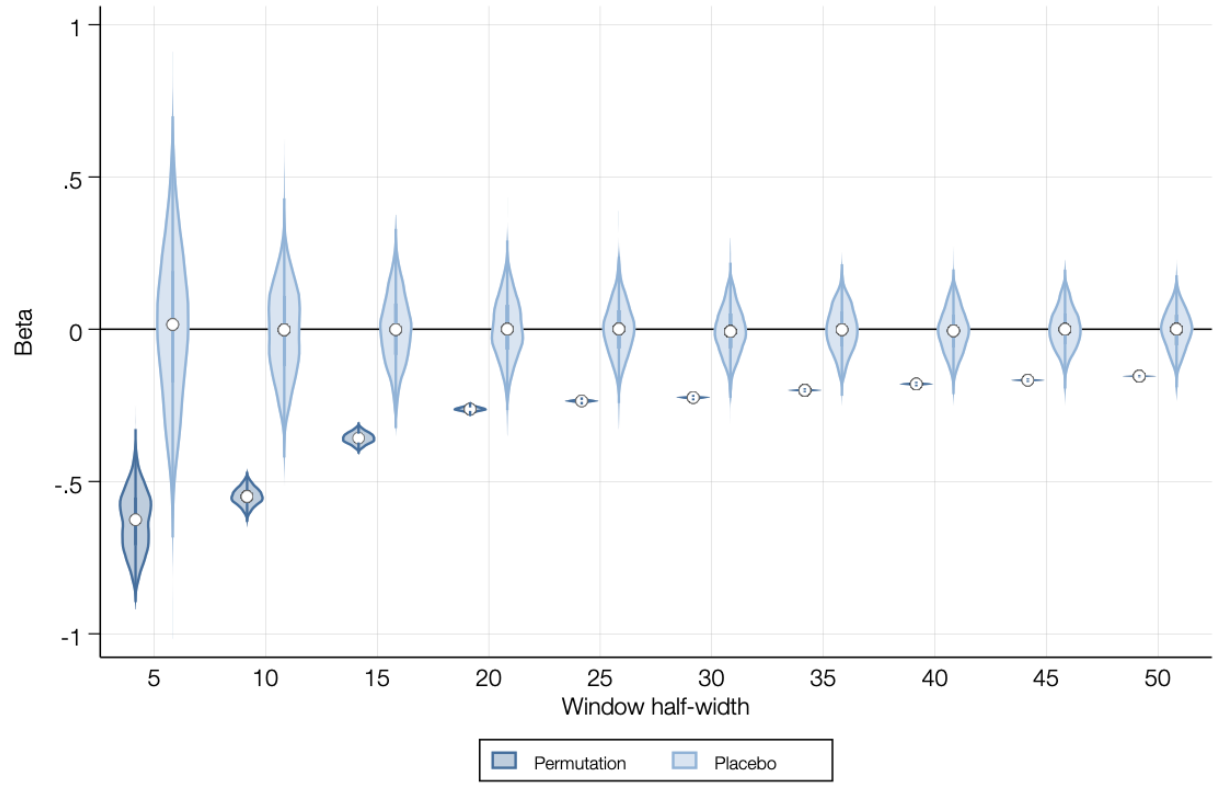
Notes: Figure shows the cumulative number of each type of medal awarded over the course of the war.

Figure A12: The estimated rate of cheapening is robust to permutations in medal order



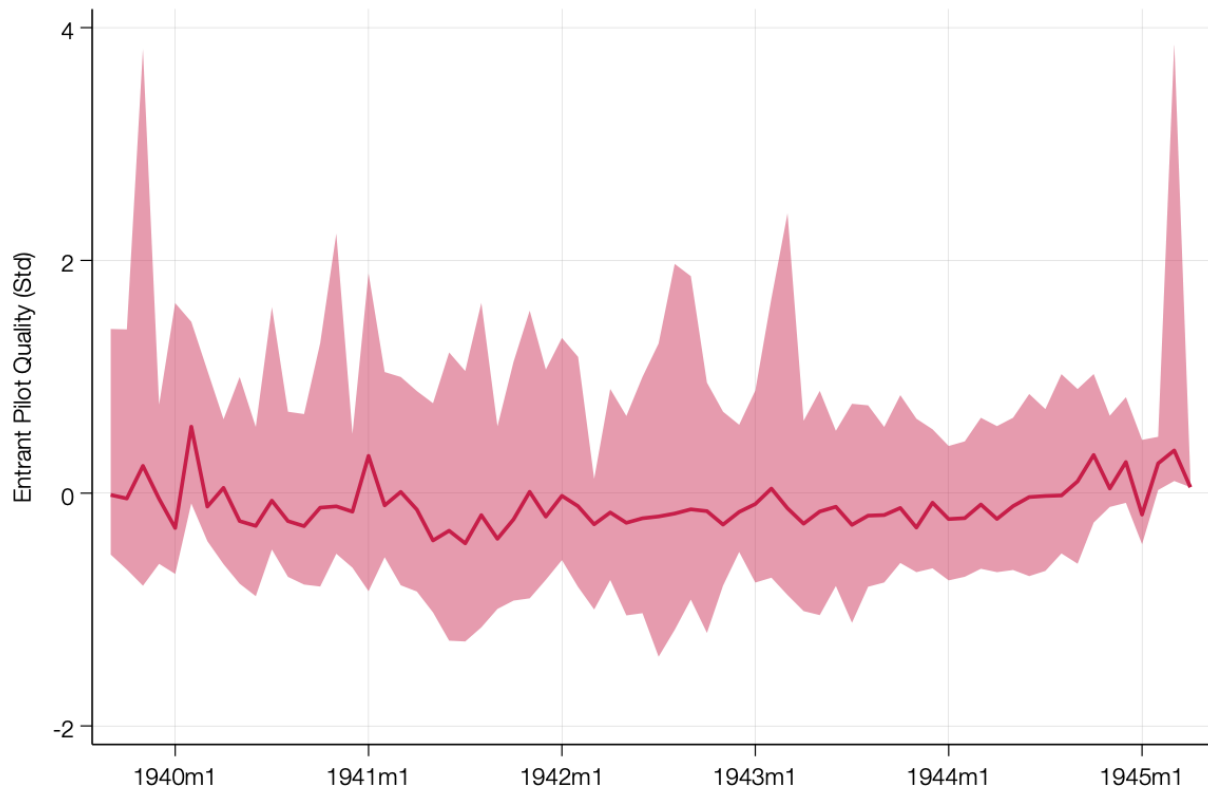
Notes: Figure shows the kernel density function of coefficient estimates of β in (4), permuted 1,000 times to resolve the ambiguous order of same-day medalists. The darker line shows the distribution of the slopes in the permutation set, where the order is permuted only *within* shared award dates. The lighter line shows the distribution of the slopes from a placebo set, where the order is permuted fully at random *across* all dates.

Figure A13: The estimated rate of cheapening is robust to the width of the comparison



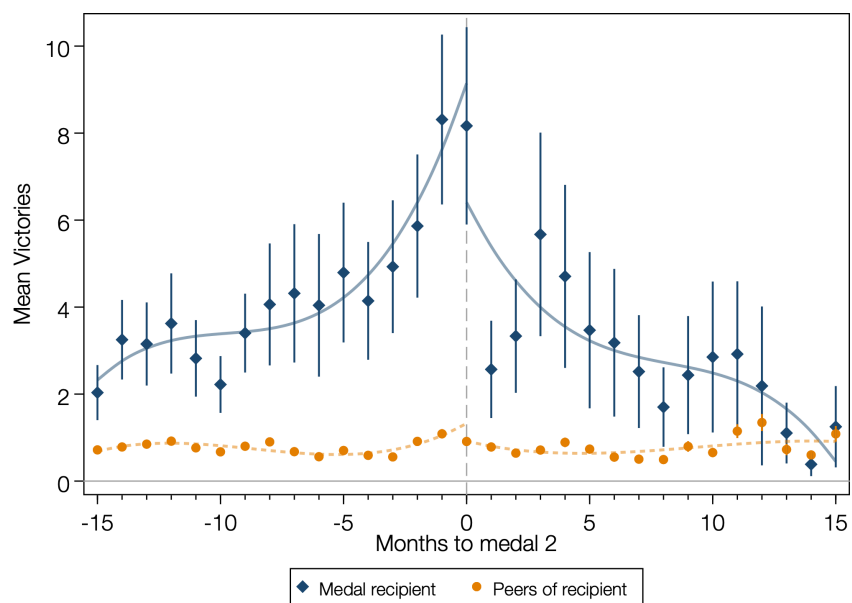
Notes: Figure shows the kernel density functions of coefficient estimates of β in (4) across a range of relative comparison window sizes W . Each distribution represents 1,000 permutations. The darker distributions show the slopes in the permutation sets, where the order is permuted only *within* shared award dates. The lighter distributions show the slopes from placebo sets, where the order is permuted fully at random *across* all dates.

Figure A14: Ability of entrant pilots over time is constant

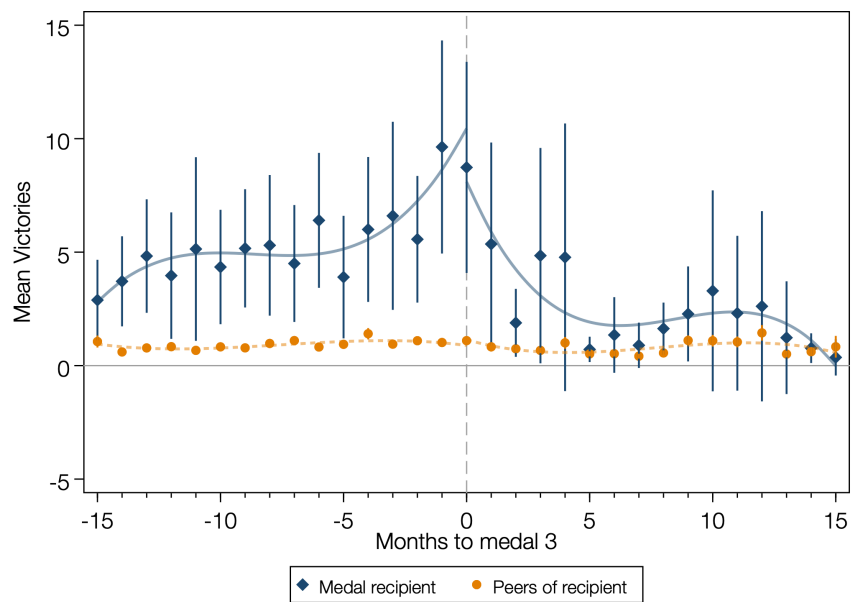


Notes: Figure shows the standardized ability ([Mas and Moretti, 2009](#)) of new pilots (in their first month as an active duty pilot) in each month of the war. The red line plots the median in each month. The shaded area spans the 10th to 90th percentiles of the data.

Figure A15: Raw victory rates over higher medal event time



(a) Medal 2



(b) Medal 3

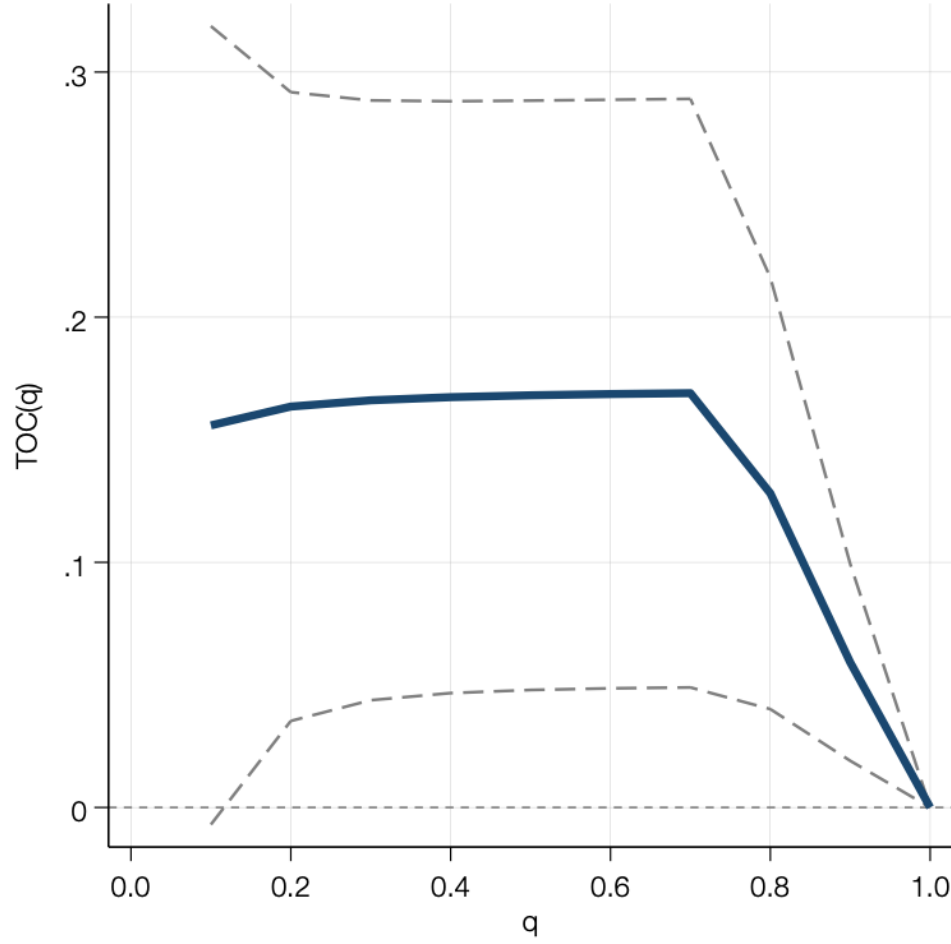
Notes: Figure shows mean monthly victories in blue over time relative to winning medal 2 (upper panel) and medal 3 (lower panel). The orange line shows the mean monthly victories of peer pilots who fly at the same airbase and in the same month as the focal pilot. Bars give the 95% confidence intervals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

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

<i>Panel A: Victories</i>				
	(1)	(2)	(3)	(4)
Inzone for medal 1	0.502*** (0.057)			0.516*** (0.057)
Inzone for medal 2		0.417*** (0.086)		0.423*** (0.086)
Inzone for medal 3			1.336** (0.629)	1.363** (0.629)
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
<i>Panel B: Victories</i>				
	(1)	(2)	(3)	(4)
Inzone for medal 1	0.220*** (0.022)			0.245*** (0.023)
Inzone for medal 2		0.199*** (0.042)		0.194*** (0.043)
Inzone for medal 3			0.025 (0.188)	0.038 (0.191)
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

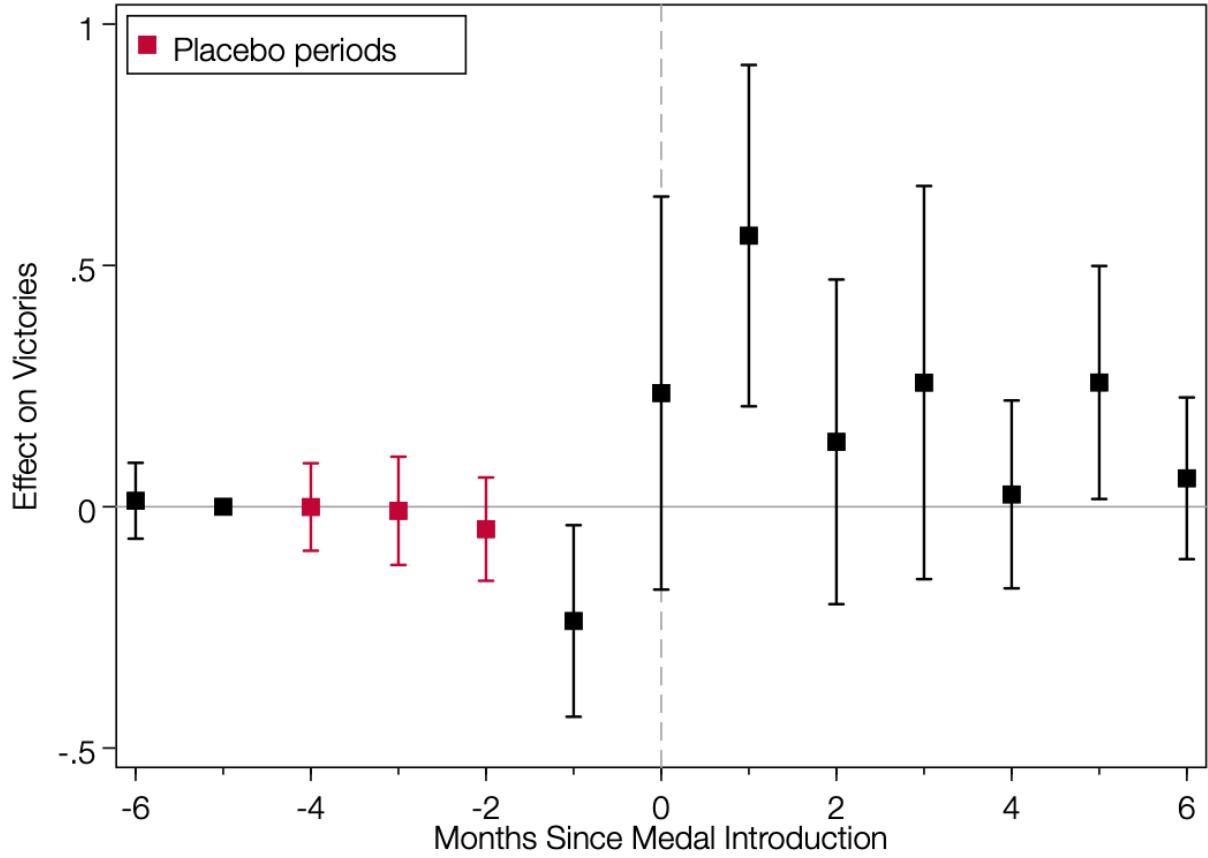
Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The table presents regression estimates. The outcome is the monthly victory rate. Inzone # indicates that the pilot's cumulative victories are no more than 15 below the quota for medal #, or at or above it, and that he does not yet hold medal #. Panel 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.

Figure A16: In zone treatment effects vary across pilots



Notes: Figure shows the targeting operator characteristic (TOC) curve for the effect of being in zone of the medal 1 quota (cumulative victories no more than fifteen below the quota, or above it, without the medal). 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. \(2025\)](#). Treatment effect sizes are ordered from largest to smallest according to the prioritization score $S(X_i)$. We then compute $\text{TOC}(q) = E[Y_i(1) - Y_i(0) | S(X_i) \geq F_{S(X_i)}^{-1}(1 - q)] - E[Y_i(1) - Y_i(0)]$. The y -axis reports the increase in the ATE over unconditional ATE for the top q proportion of units.

Figure A17: 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, $t = -5$. 95% confidence intervals from 500 bootstrap draws are indicated by the capped spikes.

Table A4: In zone response heterogeneity: prestige index and its components

	Index	Individual markers			
	(1)	(2)	(3)	(4)	(5)
Inzone=1	0.528*** (0.054)	0.509*** (0.064)	0.565*** (0.054)	0.440*** (0.049)	0.636*** (0.070)
Inzone=1 × Prestige	-0.123*** (0.028)				
Inzone=1 × Commander		-0.175** (0.074)			-0.170** (0.074)
Inzone=1 × DK Medal			-0.299*** (0.069)		-0.300*** (0.069)
Inzone=1 × Noble				-0.049 (0.184)	-0.142 (0.187)
Prestige	0.044 (0.033)				
Commander		0.366*** (0.063)			0.373*** (0.063)
DK Medal			-0.140* (0.074)		-0.155** (0.074)
Estimator	OLS	OLS	OLS	OLS	OLS
N	1,810,588	1,810,588	1,810,588	1,810,588	1,810,588
Inzone Pilots	573	573	573	573	573
FEs	Y	Y	Y	Y	Y
Outcome Mean	0.61	0.61	0.61	0.61	0.61
Outcome S.D	2.10	2.10	2.10	2.10	2.10

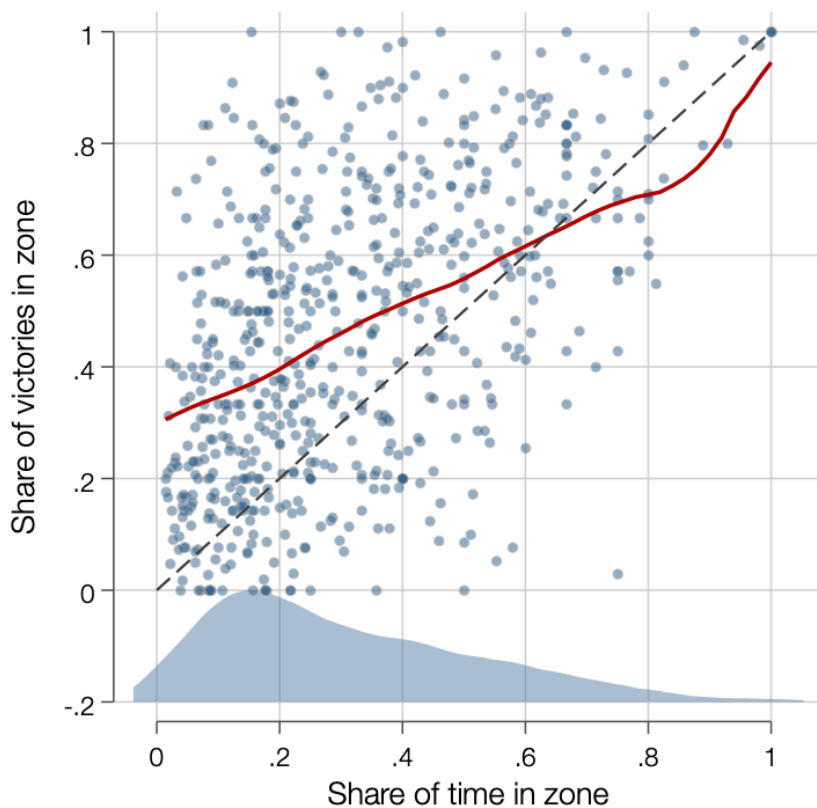
Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The table presents regression estimates. The outcome is individual monthly victories. Inzone indicates that the pilot's cumulative victories are no more than 15 below the quota for his next medal (medals 1–3), or at or above it, and that he has not yet been awarded that medal. Column 1 repeats the prestige index of [Table 2](#): the equally weighted average of standardized indicators for holding a squadron commander's rank or higher, the DK medal, and being from the nobility, rescaled to unit standard deviation. Columns 2–4 interact each marker with the in zone indicator on its own, and column 5 enters the three jointly. Nobility is time-invariant and its main effect is absorbed by the pilot fixed effects. 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.

Table A5: 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	0.493*** (0.056)	0.479*** (0.053)	0.557*** (0.066)	0.350*** (0.085)	0.348*** (0.085)	0.538*** (0.169)	1.177* (0.655)	1.182* (0.652)	1.655** (0.766)
Prestige		0.056* (0.032)	0.049 (0.034)		0.069** (0.032)	0.058* (0.031)		0.071** (0.032)	0.071** (0.032)
Inzone=1 × Prestige			-0.127*** (0.038)			-0.152 (0.093)			-0.303 (0.481)
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: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The table presents regression estimates. The outcome is individual monthly victories. Inzone indicates that the pilot's cumulative victories are no more than 15 below the quota for the medal indicated in the column header, or at or above it, and that he does not yet hold that medal. Prestige is the equally weighted index of indicators for holding a squadron commander's rank, the DK medal, and being from the nobility (each standardized, averaged, and rescaled to unit standard deviation). 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. Standard errors clustered at the level of fighter squadrons.

Figure A18: 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, which begins fifteen victories below the quota and lasts until the medal is awarded. 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

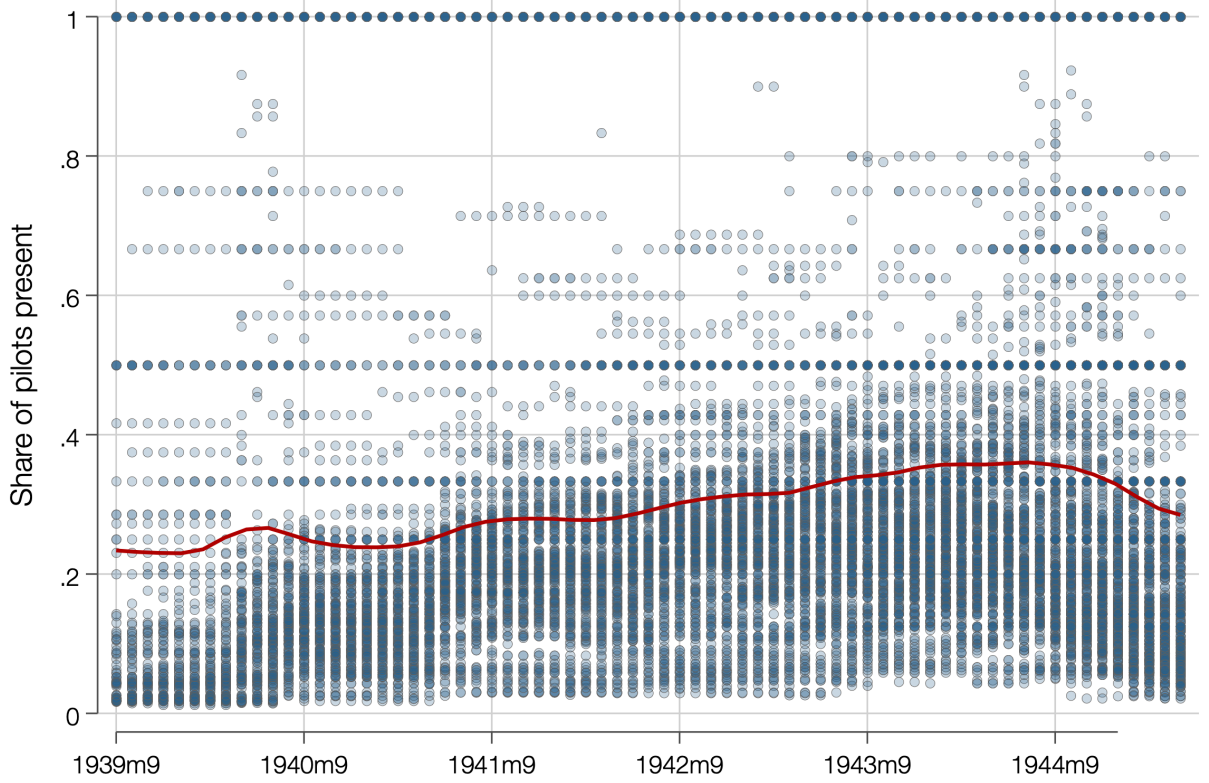
To recover time-invariant pilot ability in the presence of peer effects, we follow the empirical framework of [Mas and Moretti \(2009\)](#). Let $Y_{i,c,t}$ denote the number of verified victories achieved by pilot i in squadron c during month t . We estimate the model

$$Y_{i,c,t} = a_i + \psi_{g(c,t)} + \eta_{b(i,t)} + \gamma_{f,t} + \varepsilon_{i,c,t}, \quad (6)$$

where a_i is pilot i 's permanent productivity (ability), our object of interest. The term $\psi_{g(c,t)}$ is a fixed effect for the peer group $g(c,t)$, defined as the set of pilots serving in squadron c in month t . The fixed effect $\eta_{b(i,t)}$ captures time-invariant differences across airbases, with $b(i,t)$ the airbase at which pilot i is stationed in month t , and $\gamma_{f,t}$ captures front-by-month shocks, with f indexing combat fronts. The idiosyncratic error term is $\varepsilon_{i,c,t}$.

Identification of a_i is based on observing pilots across multiple peer compositions over time. Squadron entry was as good as randomly assigned, driven by operational needs ([Caldwell, 2012](#)). Most squadron exits are death or permanent incapacitation, so there is a large stochastic component in the composition changes. We define peers as those fellow pilots flying in the same squadron and month. As shown in [Figure A19](#), in any given month squadrons typically contain only a fraction of the total pilots who ever serve in that squadron, implying substantial within-squadron variation in peer composition over time.

Figure A19: Pilot ability estimation: squadron composition variability



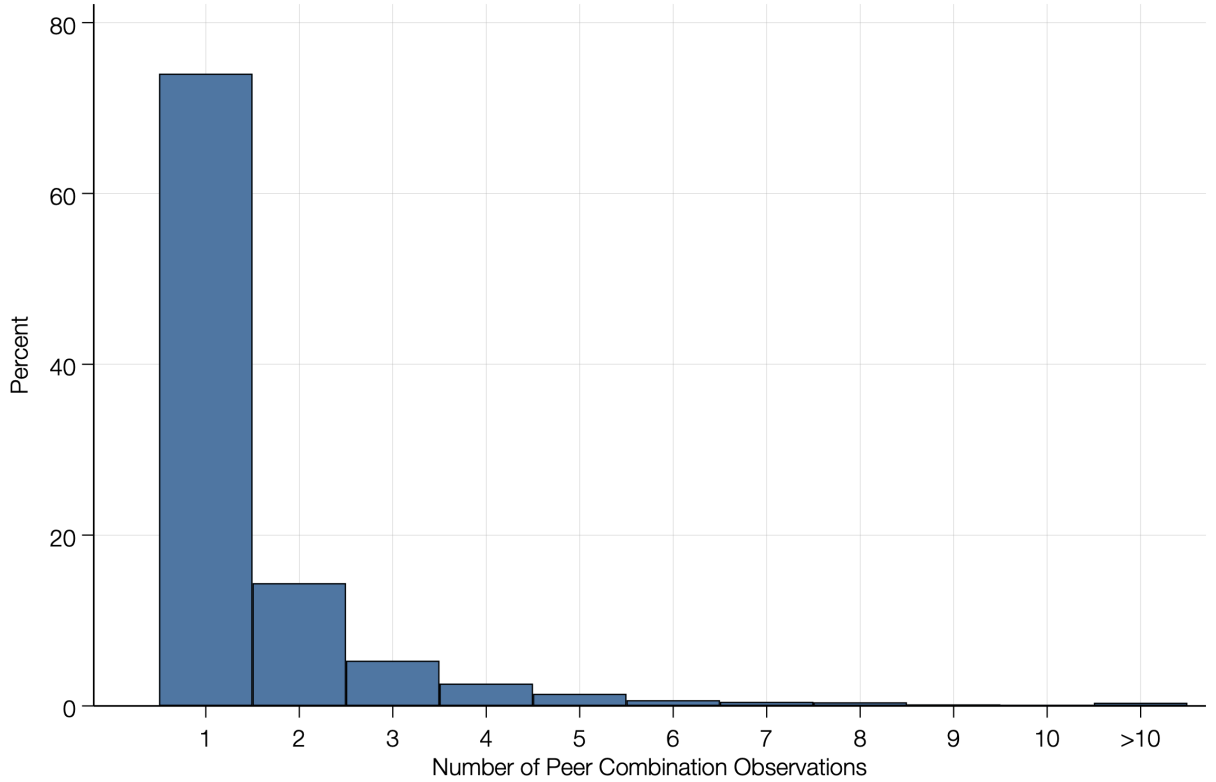
Notes: Each point is a squadron-month. The y -axis position is the number of active pilots in the squadron divided by the total number of pilots ever observed serving in this squadron. The x -axis is the monthly time. The red line is a local polynomial fit of the data.

Among 82,348 monthly pilot observations, we observe 50,075 distinct peer groups $g(c,t)$. Since the number of peer-group fixed effects is large relative to the number of observations, we cannot estimate the full set of a_i without imposing structure.

We therefore adopt the procedure from [Hamilton et al. \(2022\)](#). Specifically, we exclude peer-group fixed effects $\psi_{g(c,t)}$ that appear fewer than δ times in the panel and treat those observations as residual variation. As δ increases, we can estimate a_i for more pilots, but model fit may deteriorate if rare peer-group effects capture meaningful variation.

We examine the trade-off between model fit and coverage across a range of δ . The distribution of peer-group frequencies is shown in [Figure A20](#). Since the ratio of peer groups to total observations is high, the median peer group appears only once.

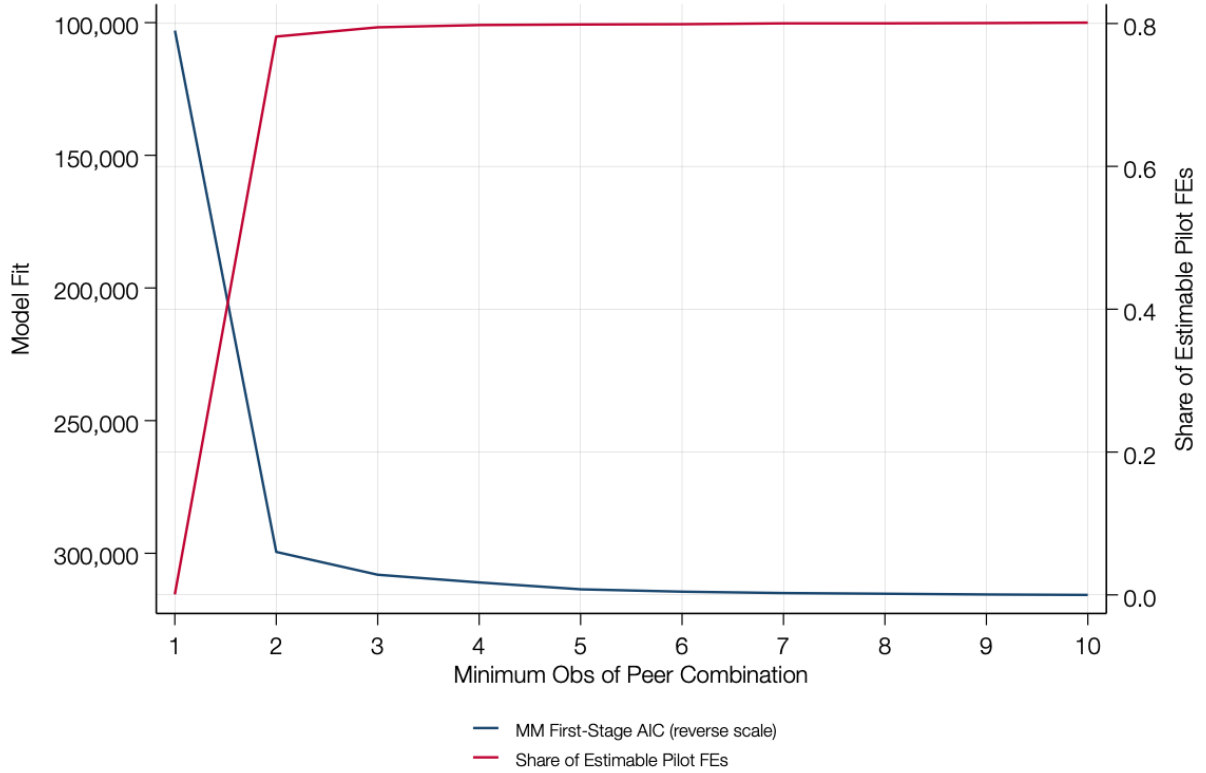
Figure A20: Pilot ability estimation: peer combination frequency



Notes: Each observation is a specific peer group $g(c,t)$. We plot the histogram showing the number of times each peer group appears in the monthly panel. The median peer group appears only once.

Model fit and estimable coverage as a function of δ are plotted in [Figure A21](#). We report the Akaike Information Criterion (AIC) for each δ , where lower values indicate better fit. We choose $\delta = 2$, the smallest value that pools the rarest peer groups; the AIC rises monotonically in δ , so this choice preserves model fit while allowing identification of permanent productivity for the majority of pilots.

Figure A21: 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 minimum number of observations required for a peer group $g(c,t)$ to receive its own fixed effect. We successively exclude peer-group fixed effects observed fewer than δ times and refit the model. The AIC is reported for each specification. With a stronger δ restriction, we can estimate a greater number of individual pilot permanent productivities.

In [Table A6](#), we compare the observable characteristics of pilots whose a_i is estimable under this choice of δ to those whose permanent productivity cannot be recovered. As expected, estimable pilots serve longer and change squadrons more frequently.

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

	Estimable	Non-Estimable	T-test Difference
Mean Monthly Victories	0.787 (1.050)	1.106 (0.853)	-0.319***
Mean Front	0.343 (0.437)	0.295 (0.454)	0.048***
KCR	0.102 (0.303)	0.011 (0.104)	0.091***
Total Victories	13.191 (25.426)	2.201 (3.932)	10.990***
Tenure	19.597 (17.866)	4.067 (8.641)	15.530***
# Squad Changes	0.775 (1.309)	0.096 (0.522)	0.679***
N	3,972	1,109	5,081

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Characteristics of pilots for whom we can and cannot estimate individual permanent productivity a_i using the modified [Mas and Moretti \(2009\)](#) procedure with peer-group fixed effects restricted to appear at least $\delta = 2$ times. T-tests report differences in means between groups.

C Monte Carlo Validation of the Medal Anticipation Effects Estimator

Here we describe the details of the Monte Carlo exercise used to validate our estimator for recovering medal anticipation effects, as introduced in [Section 2](#).

Data-generating process. We simulate a panel with $N = 400$ pilots observed over $T = 70$ monthly periods. Each pilot draws a time-invariant ability parameter

$$a_i \sim \log \mathcal{N}(0, 0.4^2),$$

and faces a fixed quota $q_i = 20$.

Let cumulative victories be

$$V_{it} = \sum_{\tau \leq t} Y_{i\tau}.$$

Baseline victories follow a Poisson process with a conditional mean

$$\mathbb{E}[Y_{it}^{\text{base}} \mid a_i] = \exp(a_i).$$

We draw

$$Y_{it}^{\text{base}} \sim \text{Poisson}(\exp(a_i)).$$

Medal anticipation effects. We define distance to quota as

$$d_{it} = V_{it-1} - q_i,$$

so that $d_{it} \leq 0$ corresponds to being below the quota.

Anticipatory effort is added only while the pilot remains below the quota:

$$Y_{it}^{\text{extra}} = \begin{cases} e(d_{it}) & \text{if } d_{it} \leq 0, \\ 0 & \text{if } d_{it} > 0, \end{cases} \quad (7)$$

where

$$e(d_{it}) = \frac{\alpha}{[1 + \exp(\beta(d_{it} - c))](1 + \Gamma a_i)}.$$

The parameters $\{\alpha, \beta, c\}$ are chosen to approximate the empirical shape in [Figure 2](#). The parameter $\Gamma \geq 0$ allows the magnitude of anticipation effects to vary weakly with ability.

Total victories are

$$Y_{it} = Y_{it}^{\text{base}} + Y_{it}^{\text{extra}}.$$

Under the null hypothesis of no anticipation effect, we set $\alpha = 0$.

The event time is defined as the stopping time

$$E_i = \inf\{t : V_{it} \geq q_i\}.$$

Exit process. To incorporate survivorship selection, we introduce endogenous exit. Let standardized victories be

$$y_{it} = \frac{Y_{it} - \bar{Y}_t}{\sigma_t},$$

where $\bar{Y}_t$ and σ_t are the cross-sectional mean and standard deviation of Y_{it} in month t .

Risk is defined as

$$r_{it} = \Sigma y_{it} (1 + \Phi \cdot (1 - a_i)) - \Omega a_i + \nu_{it},$$

where $\nu_{it} \sim \mathcal{N}(0,1)$.

Parameters satisfy $\Sigma \geq 0$, $\Phi \geq 0$, and $\Omega \geq 0$. Higher victories increase risk, while higher ability reduces baseline risk.

In each period, pilots above the $(1 - \kappa)$ quantile of the risk distribution exit permanently. We set $\kappa = 0.05$, so that 5% of pilots exit each month.

Placebo quotas. For each pilot i , we draw $P = 25$ placebo quotas

$$q_i^{(p)} \sim \text{Unif}[\underline{\mu}q_i, \bar{\mu}q_i], \quad p = 1, \dots, P,$$

with $\underline{\mu} = 0.1$ and $\bar{\mu} = 2.0$.

For each placebo quota we define a placebo medal time

$$E_i^{(p)} = \inf\{t : V_{it} \geq q_i^{(p)}\}.$$

[Figure A5](#) reports the average event-study coefficients across 50 Monte Carlo repetitions. In each repetition, we generate $P = 25$ placebo quotas per pilot and estimate both the naive and placebo-adjusted event-study specifications.

The dashed gray line shows the naive estimator $\{\hat{\theta}_j^{\text{naive}}\}$ obtained from (2). Even under the null, it produces a pronounced pre-quota ramp and a spurious positive post-quota effect due to the survivorship effect.

The solid red line shows the adjusted estimator $\{\hat{\theta}_j^{\text{adj}}\}$ from (3). Across both the null and alternative data-generating processes, the adjusted estimator closely tracks the true effect θ_j , and its confidence band covers the truth at all horizons except the crossing month under the alternative, where it recovers about half of the true effect.

The simulation, therefore, confirms that the placebo-threshold procedure reliably removes the mechanical bias b_j .

D Complementarity in Squadron Output

In the main text, we treat a pilot’s output as determined by his own skill, equipment and motivation, and by aggregate combat conditions. Squadron composition can shift a pilot’s level of output—our ability measure nets it out (Appendix B)—but we assume that pilots’ abilities do not interact: two able pilots produce no more together than the sum of what each would produce alone. Yet some joint production certainly occurred. Wingmen mattered—German fighters flew as pairs of leader and wingman—and the Luftwaffe pioneered tactics such as the “four-finger” formation, in which two such pairs fought as a single unit. Rather than assume that such effects were negligible, this appendix asks whether they shaped outcomes: did able pilots make one another more productive? We use the Shapley value, designed for dividing the output of a team among its members (Shapley, 1953). We find no evidence that they did: complementarity in squadron output is small and statistically insignificant.

Method. We treat each squadron in each calendar month as a cooperative game whose players are the pilots on its roster. Our panel contains only pilots with at least one recorded victory, most of whom enter it in the month of their first victory. In our estimation sample, squadron-months with at least two pilots yield 9,735 games, 73,435 pilot-month observations and 4,696 pilots, September 1939 to May 1945. The median roster has 6 pilots, the largest 26.

Let r_i denote pilot i ’s victory rate *before* the month—victories per month in our panel up to the end of the previous month, so that it contains none of the victories it is meant to explain. The classic Shapley value requires the output of every possible sub-team. We never observe sub-teams—only the full roster of each squadron in each month. We therefore estimate a value function, which gives the output of any coalition, from how squadron output varies across rosters, and apply the Shapley formula to the estimate. Let n_S be the number of pilots in a coalition S . Its value is

$$v(S) = \sum_{i \in S} g(r_i) + h(n_S) + \kappa \sum_{i < j \in S} r_i r_j. \quad (8)$$

The first term is each pilot’s own contribution, an arbitrary function g of his own rate. The second is an arbitrary function h of squadron size. Both are additive: they let a pilot’s output rise faster or slower than his rate, and let large squadrons score more or less per head, without any pilot affecting another. The third term is complementarity: with $\kappa > 0$, two able pilots together produce more than the sum of their separate contributions. The

Shapley value then has a closed form,

$$\phi_i = \underbrace{g(r_i) + h(n_S)/n_S}_{\text{stand-alone}} + \underbrace{\frac{\kappa}{2} r_i \sum_{j \neq i} r_j}_{\text{complementarity}} , \quad (9)$$

since each pair's joint term is split equally between its two members. If $\kappa = 0$, a pilot's Shapley value equals his stand-alone contribution: there is no joint surplus to share. Testing whether Shapley values exceed stand-alone contributions is therefore the same as testing $\kappa = 0$.

Estimation. We estimate v on the observed games, regressing squadron victories in a month on the pairwise term, the number of members in each own-rate group (pilots with a zero rate, plus groups of positive rates cut at their 5 percent quantiles; this traces out g), one indicator per squadron size (which traces out h), and squadron, airbase and front-by-month fixed effects. Flexible g and h matter. The pairwise term rises with the square of squadron size and of total ability; with linear g and h , any curvature in those additive terms would be loaded onto κ and mistaken for complementarity. The fixed effects mean that κ is identified from changes in a squadron's roster over time, net of where and when it fought.

Results. [Figure A22a](#) shows the pattern in the data. Net of the flexible additive terms and the fixed effects, squadron victories show no tendency to rise with the pairwise products of members' rates. [Table A7](#) reports the estimates. With g and h flexible (column 3), $\kappa = -0.0003$ (s.e. 0.0103). Coarser own-rate groups (column 4) give 0.0026 (s.e. 0.0103), and linear g and h (column 1) give 0.0023 (s.e. 0.0111). None is close to significant.

To judge magnitudes, we sum the complementarity part of (9) over all pilot-months and divide by total victories. Complementarity accounts for -0.1% of all victories at the point estimate of column 3, and for 8.7% at the upper end of its 95% confidence interval. Spillovers from aces should be largest among the ablest pilots. Among the top tenth of positive rates¹, the upper value is 11.8% of their own victories. [Figure A22b](#) compares the complementarity part of the Shapley value with what pilots actually scored, decile by decile of their rate.

Summary. The Shapley decomposition offers no evidence of complementarity among pilots: the complementarity term is small and insignificant once each pilot's own contribution and squadron size enter flexibly. Two limits apply. Rosters contain only pilots who score at least once in the war, so the test concerns complementarity among scorers. And because the test concerns total squadron output, it cannot detect a reallocation of credit within a

¹We restrict to pilots in at least their sixth month in the panel, whose rates no longer mainly reflect the month of their entry.

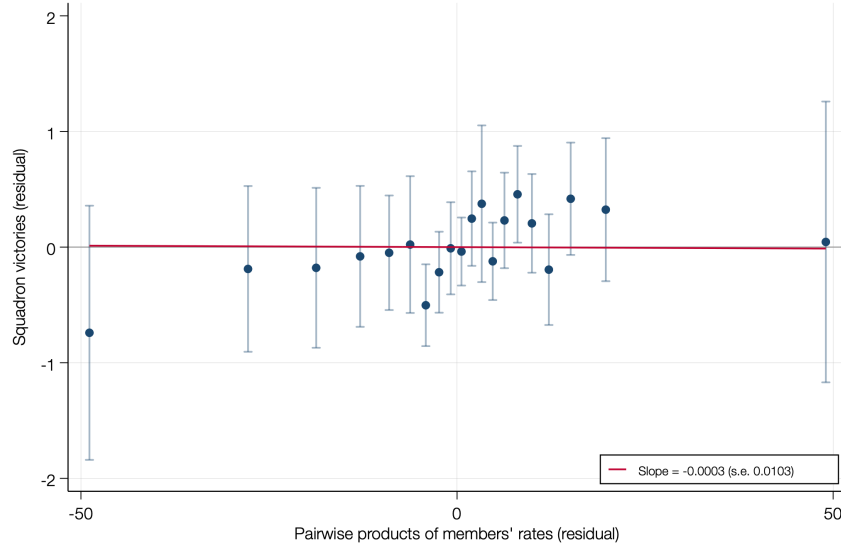
squadron, such as a leader being credited with victories that he and his wingman achieved together. Section 2.2 addresses that concern directly for pilots near the quota.

Table A7: Coalition value function: complementarity in squadron output

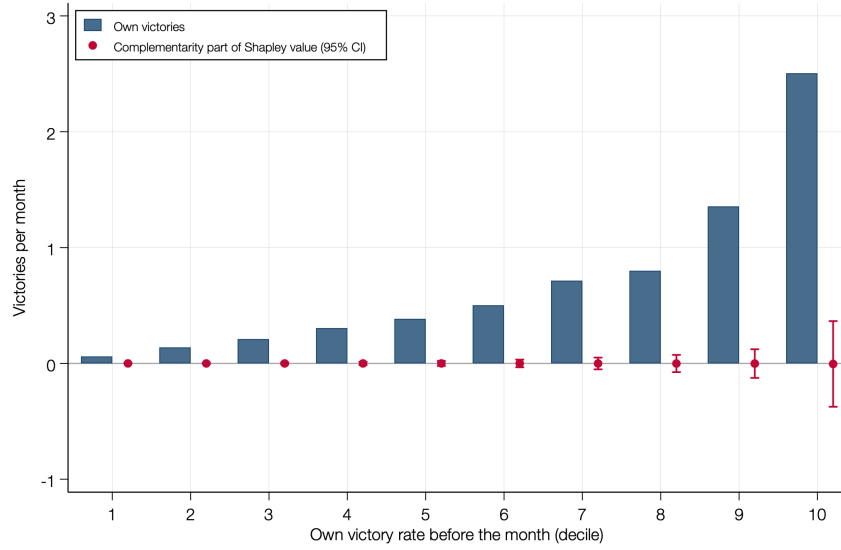
	Squadron victories			
	(1)	(2)	(3)	(4)
Pairwise products of rates	0.0023 (0.0111)	-0.0058 (0.0112)	-0.0003 (0.0103)	0.0026 (0.0103)
Sum of members' rates	0.3925*** (0.1172)	0.4718*** (0.1244)		
Pilots on roster	0.8075*** (0.0964)			
Own-rate term g	Linear	Linear	Flexible (5% groups)	Flexible (10% groups)
Roster-size term h	Linear	Flexible	Flexible	Flexible
Outcome Mean	5.29	5.29	5.29	5.29
Squadron-months	9,735	9,735	9,735	9,735
R ²	0.591	0.595	0.616	0.615

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. OLS estimates of (8), one observation per squadron and calendar month with at least two pilots, September 1939 to May 1945. The outcome is squadron victories in the month. "Pairwise products of rates" is $\sum_{i < j} r_i r_j$, with r_i the pilot's victories per month in the panel up to the end of the previous month; its coefficient κ measures complementarity. Own-rate term g : linear (sum of members' rates) or flexible (number of members in each own-rate group; pilots with a zero rate form the base group, and positive rates are cut at the 5 or 10 percent quantiles of positive pilot-month rates). Roster-size term h : linear (number of pilots) or flexible (one indicator per size, sizes above 20 pooled). All columns include airbase, front-by-month and squadron fixed effects, with each squadron-month assigned the most common base and front of its members; squadron-months in which no pilot has a recorded base, and singletons within a fixed-effect group, are dropped. Standard errors two-way clustered by squadron and calendar month in parentheses.

Figure A22: Shapley decomposition of squadron output



(a) Squadron output and pairwise products of members' rates



(b) Complementarity part of the Shapley value and own victories

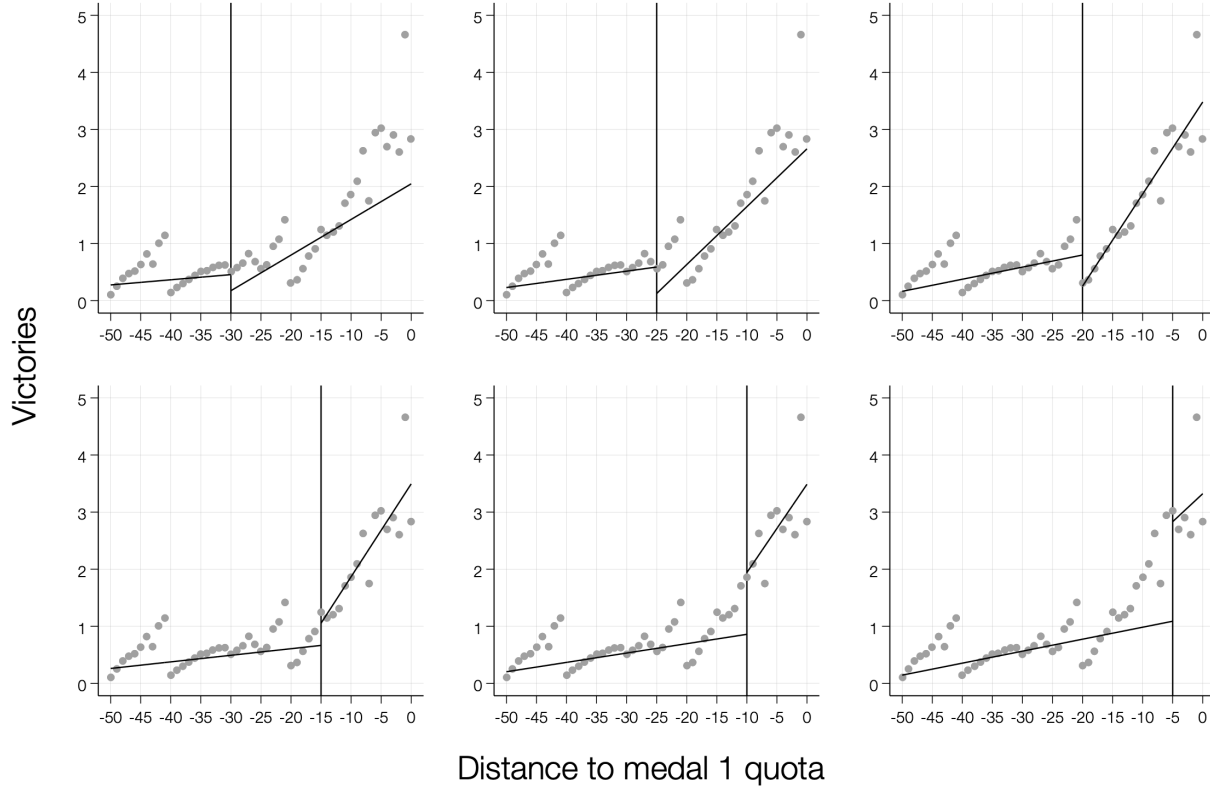
Notes: Panel (a) is a binned scatter of squadron victories in a month against the sum of pairwise products of members' victory rates before the month, both net of the controls and fixed effects of [Table A7](#), column 3. The 9,735 squadron-months are divided into twenty equal-sized bins; capped spikes are 95% confidence intervals of each bin mean, clustered by squadron. The line has the regression slope, $\kappa = -0.0003$ (two-way clustered s.e. 0.0103). Panel (b) shows, by decile of the pilot's victory rate before the month among pilot-months with a positive rate, mean own victories per month (bars) and the mean complementarity part of the Shapley value, $\frac{\kappa}{2} r_i \sum_{j \neq i} r_j$, at the point estimate of column 3 (dots), with its range when κ runs over its 95% confidence interval, $[-0.021, 0.020]$ (capped spikes).

E 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 having cumulative victories no lower than $q_{it} + D$ —at most $|D|$ below the quota q_{it} , or already above it—while not yet holding the medal. A pilot who meets the quota thus stays in zone until the medal is awarded: the quota was an informal criterion, awards followed with a lag, and some pilots who met it never received the medal, so the incentive is resolved only by the award itself. We calibrate D by estimating regression kink analysis around a series of candidate entry points into the zone, $D \in \{-30, -25, \dots, -5\}$. We show in Figure A23 that the kink in the victory rate—the change in its slope with respect to quota distance, estimated using Calonico et al. (2015)—is most prominent at $D = -15$.

Figure A23: Local polynomial estimates of zone entry

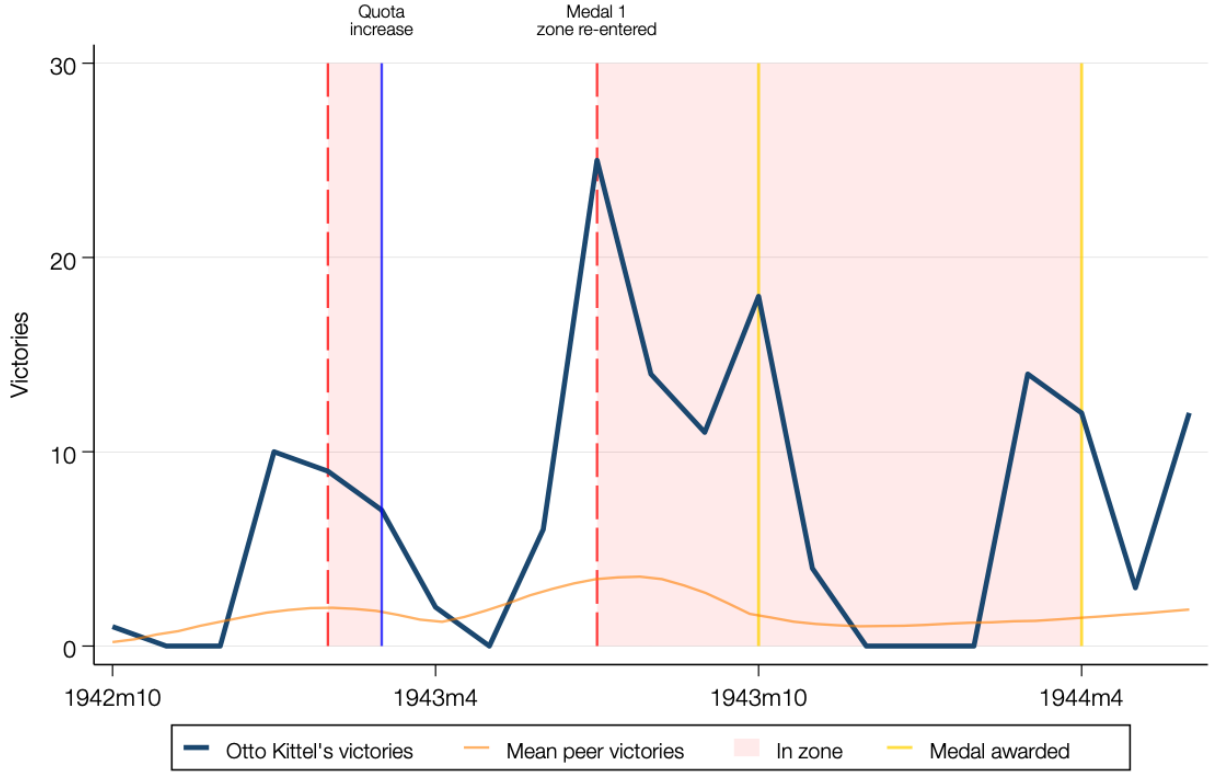


Notes: Figure shows a series of regression discontinuity plots estimated around candidate “in zone” entry points, the stock of pilot victories is distance $D \in \{-30, -25, \dots, -5\}$ below the informal victories quota for medal 1. Using the method of Calonico et al. (2015), the data is partitioned 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: $\mathbb{1}(q_{it} - q_{i,t-1} > 0)$. When the quota is increased, pilots are moved out of the zone even as their victory tally is constant. Under our definition, an increase can only move pilots out of the zone, never into it. Given the relative abundance of quota increases (relative to decreases) in our data, we expect to be well-powered to detect effects (see [Figure A1](#)). The first-stage and reduced-form regressions should yield a negative coefficient on the instrument: quota increases push the marginal pilot *away* from the medal zone.

We illustrate the causal mechanism described above with the case of Otto Kittel ([Figure A24](#)). 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. In July he scores 25 victories, entering the zone and passing the quota in a single month, and he continues to perform strongly in August and September while the award is pending. Finally, he receives medal 1 in October, by which time his tally already exceeds the quota for the next medal. We see that in the period after the quota increase, performance remains low, since the medal quota is distant. As 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 A24: Quota changes and performance: case study evidence



Notes: Figure shows the performance of fighter ace Otto Kittel (b. 1917) over time. The dark blue line gives the monthly victory score. The orange line gives the mean score of squadron peers. Dashed red vertical lines mark the pilot's entry into the zone of a medal quota (fifteen victories below it); shaded periods are in zone, from entry until the medal is awarded or a quota increase moves him out of the zone. The gold lines indicate when a medal was awarded. The vertical blue line 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_{it} and the in zone treatment $\mathbb{1}(V_{it} - q_{it} \geq D) \cdot \mathbb{1}(M_{it} = 0)$, where our cutoff is $D = -15$, $V_{it} = \sum_{\tau \leq t} Y_{i\tau}$ is cumulative victories, and M_{it} indicates that pilot i holds the medal in period t . The exclusion restriction requires that quota increases affect pilot performance only by altering proximity to the medal quota. Conditional on fixed effects for period and front, remaining shocks to combat conditions that correlate with quota increases are likely to be small. Quota increases are motivated by longer-run prestige concerns regarding the supply of medals, rather than by short-run battlefield conditions.

One might be concerned that the same quota increases affect multiple pilots who compete with their peers to achieve the medal first (Ager et al., 2022). First, variation in q_{it} will affect only pilots without the medal whose victories V_{it} are close to or above q_{it} . Figure A3 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 in zone pilot-months in

which a squadron peer is also in zone, to preclude our estimation from capturing competitive rather than motivational responses to the quota increase. 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 increase 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 increase spills over across pilots (e.g., on the same airbase).

In [Table A2](#), 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 report 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 quota increases 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 on in zone peers. In columns 5 and 6, we drop in zone pilot-months with an in zone squadron peer, to remove the potential for competitive responses. The final column shows the preferred specification, with covariates added 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 at conventional levels in all three 2SLS specifications. Overall, we find effects that are consistent with, though larger than, our preferred event-study approach.

F Alternative Ability Estimation in Cheapening Exercise

In this section, we address the concern from [Section 3.2](#) that our estimates of medal cheapening over time are biased by the construction of the ability measure. The concern is that many victories—which inform the ability estimate—are attained while under the influence of the medal incentives. The estimates of cheapening might then reflect a circularity: pilots who are more receptive to the medal incentives are measured as higher ability, when their out-performance is instead explained by effort.

Our approach is to demonstrate the robustness of our estimates to using a measure of ability that excludes periods when pilots are in zone for a medal quota. First, we re-estimate our [Mas and Moretti \(2009\)](#) ability measure on the remaining 75,758 of the 82,348 panel observations. The resulting measure has a Spearman rank correlation of 0.881 with the main ability estimate at the pilot level (0.791 among pilots who win medal 1).

Second, we re-estimate the cheapening coefficient β as in Section 3.2, using the same 1,000 permutations of same-day awards. The coefficient $\hat{\beta}_{excl}$ is presented in Table A8 and is substantially attenuated relative to the preferred estimate, which we reproduce in the first row for convenience. Third, we compare the new estimate to the expected β under random sub-sampling of the panel: for each pilot, we drop at random as many months as he spends in zone. This step is essential because in-zone months carry a disproportionate share of medalists' victories (Figure A18); excluding them makes the ability estimates noisier, which compresses systematic rank differences and mechanically shrinks the estimated slope irrespective of the circularity concern. We report the mean coefficient $\bar{\beta}_{random}$ across 50 replications, which sits between the preferred and excluded estimates and remains statistically significant (one-sided $p = 0.0472$).

Finally, to separate attenuation from circularity, we compute in the last row the attenuation-corrected coefficient, $\hat{\beta}_{clean} = \hat{\beta}_{full} \hat{\beta}_{excl} / \bar{\beta}_{random}$ —the cheapening slope implied for an ability ranking purged of in-zone months, rescaled to full-panel precision through the attenuation factor $\bar{\beta}_{random} / \hat{\beta}_{full}$.

Our inference procedure is as follows. Each medalist enters the stacked regression both as a focal pilot and as a neighbor of other medalists, and the windows overlap, so standard cluster bootstraps over medalists misstate the sampling variability of the slope. We therefore hold the award order fixed and use a wild bootstrap. The estimator is linear in the ranks, and in each of 10,000 replications we perturb every medalist's deviation from the mean rank of his window neighbors with a Rademacher multiplier. The multiplier is common to all specifications, so the replications preserve the correlation between the rows. Each replication also resamples the 50 random-drop replications with replacement, so that the simulation noise in $\bar{\beta}_{random}$ carries into $\hat{\beta}_{clean}$. Estimated rankings and tie-breaking permutations are held fixed. We report 95% percentile intervals, and one-sided p -values are the share of replications with a non-negative coefficient. For the first three rows, the randomization p -values of Section 3.2 are similar (0.0050, 0.1950, and 0.0510).

The implied circularity component of the full-panel coefficient is $\hat{\beta}_{full} - \hat{\beta}_{clean} = -0.0928$ (95% interval $[-0.4605, 0.1550]$; one-sided $p = 0.1389$), which is not statistically distinguishable from zero. Random drops reproduce 66% of the coefficient decline under the in-zone exclusion, and the attenuation-corrected coefficient retains 61% of the magnitude of the headline estimate. Its interval is wide because the attenuation factor is itself estimated with noise.

Table A8: Cheapening slope: attenuation check for the exclusion of in-zone months

Sample	Coefficient	95% interval	One-sided p	Pilot-months	Medalists ranked
Full panel	-0.2351	[-0.4046, -0.0725]	0.0020	75,446	402
Excluding in-zone months	-0.0805	[-0.2648, 0.1000]	0.1958	69,118	401
Random count-matched drops	-0.1330	[-0.2935, 0.0209]	0.0472	69,091	401
Attenuation-corrected coefficient	-0.1423	[-0.4353, 0.3377]	0.1530		

Notes: Table reports the cheapening coefficient of Section 3.2—the slope of relative ability rank on relative medal order among the 412 recipients of medal 1, with a window half-width of 25, stacked over 1,000 permutations of same-day awards—when the [Mas and Moretti \(2009\)](#) ability measure is estimated on alternative samples. The first row uses the full monthly panel and reproduces the main-text estimate. The second row excludes all pilot-months in which a pilot is in zone for medals 1–3. The third row instead drops, for each pilot, a randomly chosen set of months equal in number to his in-zone months, holding the sample size of the second row fixed while breaking the link to in-zone status; it reports the mean coefficient across 50 replications (standard deviation across replications 0.0408). The final row reports the attenuation-corrected coefficient, $\hat{\beta}_{clean} = \hat{\beta}_{full} \hat{\beta}_{excl} / \bar{\beta}_{random}$. Intervals are 95% percentile intervals from 10,000 wild bootstrap replications that hold the award order and estimated rankings fixed, perturb each medalist’s rank deviation with a multiplier common to all rows, and resample the random-drop replications; one-sided p -values are the share of replications with a non-negative coefficient (null: no cheapening). Pilot-months is the number of observations entering the ability estimation after the sample restriction and dropping of singletons; Medalists ranked is the number of medal-1 recipients with an estimable ability measure (mean across replications in the third row, rounded to the nearest integer). Pilot fixed effects are absorbed, so pilots reduced to singleton observations are not estimable; peer-group fixed effects are restricted to appear at least $\delta = 2$ times.