

---

# Släpp dataseten fria!

*"You can fool some people sometimes,  
but you can't fool all the people all the time"*

Bob Marley

Den naturvetenskapliga forskningen drabbades nyligen av en rejäl chock. Det visade sig nämligen att den världsberömda och Nobelpristippade fysikern Jan Hendrik Schön både hade använt fabricerade data och framställt faktiska data på ett oriktigt sätt. Detta i ett flertal studier publicerade i framstående vetenskapliga tidskrifter som *Science* och *Nature*. Fusket uppdagades när forskarkollegor försökte replikera Schöns resultat utan att lyckas. Detta visar om något att replikationsstudier är av högsta vikt för att upprätthålla trovärdigheten inom forskningen. Replikationsstudier är nämligen det bästa sättet att kontrollera publicerade resultat och de ger också en möjlighet att undersöka resultatens känslighet med avseende på val av modell och metod.

Att det är viktigt med replikationsstudier är dock inget nytt, utan ett faktum som är väl känt i hela den internationella forskarvärlden. Detta gestaltar sig t ex genom att många av de ledande internationella tidskrifterna kräver att man lämnar in sitt dataset då man publicerar en artikel. Från det ekonomiska området kan t ex nämnas *Journal of Applied Econometrics* ("A special feature of the Journal is its emphasis on the replicability of results by other researchers. To achieve this aim, authors are expected to make available a complete set of the data

used as well as any specialised computer programs employed through a readily accessible medium ...", citat från tidskriftens hemsida) och *Journal of Business and Economic Statistics* (JBES) ("The journal requires that data source be clearly indicated in the article; furthermore, upon acceptance of papers for publication, authors are asked to make data available for downloading at the JBES FTP archive", citat från tidskriftens hemsida). Till följd av det uppdagade fusket med framför allt Schöns forskning pågår det för närvarande dessutom en intensiv debatt inom naturvetenskapen om hur man skall öka tillgängligheten av data som använts i forskningsrapporter.

Även vissa statliga myndigheter i Sverige har anammat policyn att tillhandahålla de data som använts i publicerade rapporter och utredningar; ett exempel är Institutet för arbetsmarknadspolitisk utvärdering (IFAU). I skenet av den svenska offentlighetsprincipen, och med tanke på att det är medborgarnas pengar som används, förhoppningsvis till medborgarnas nytta, borde denna policy vara given för alla typer av offentliga utredningar och forskningsrapporter. Gång efter annan rapporteras dock om fall där forskare och/eller myndigheter vägrar lämna ut sina dataset till kollegor för replikering. Ett aktuellt exempel från

Sverige och den samhällsvetenskapliga sidan är Sören Wibes misslyckade försök med att få data då han ville replikera resultaten i Fredrik Bergströms och Mikael Sandströms ESO-rapport om konkurrens mellan privata och kommunala skolor. Undertecknade har dessutom själva för några år sedan stött på samma problem i samband med en ESO-rapport om kommunal ekonomi skriven av Stefan Fölster; inte heller då ville vare sig forskaren själv eller ESO lämna ut data. Det är naturligtvis olyckligt att andra forskare inte ges möjligheten att replikera studierna. Detta sänker resultatens trovärdighet och väcker i värsta fall misstankar om fusk. Särskilt allvarligt är det om resultaten får stort genomslag i media.

Vi anser att fri och enkel tillgång till dataset som använts i forskningsrapporter av olika slag bör vara den grund som all seriös empirisk forskning vilar på. Denna uppfattning verkar dock alltså inte delas av alla i

Sverige, inte ens av offentliga organ vars uppgift är att försöka få fram "sanningen" inom den offentliga sektorn. ESO läggs nu ned. Men det kommer en motsvarighet med inriktning på miljöforskning och det kommer att uppstå liknande institutioner i framtiden. Vi vill råda alla organisationer som publicerar forskningsrapporter eller utredningar att de tillhåller de använda dataseten för att säkra kvaliteten och tillförlitligheten i sina resultat. Det bör vara en självklarhet att de dataset som används vid studier som ESO:s skall göras tillgängliga för den som så önskar, helst på ett enkelt sätt via någon hemsida på internet.

Vi vill därför avsluta våra två år som redaktörer för *Ekonomisk Debatt* med en uppmaning till såväl statliga myndigheter som forskningsinstitutioner: Släpp dataseten fria!

*MATZ DAHLBERG  
EVA JOHANSSON*